

**INSTITUTO SUPERIOR DE ECONOMIA E GESTÃO
UNIVERSIDADE TÉCNICA DE LISBOA**

MESTRADO EM FINANÇAS

**AVALIAÇÃO DO RISCO DE CRÉDITO DE CLIENTES EMPRESARIAIS
LEVANTAMENTO DE REQUISITOS E ESTIMAÇÃO DE MODELOS**

VERA CRISTINA SANTIAGO CAEIRO

Júri:

Presidente: Doutora Raquel Maria Medeiros Gaspar, professora associada do Instituto Superior de Economia e Gestão da Universidade Técnica de Lisboa;

Vogais: - Doutor Pedro Luís Pereira Verga Matos, professor auxiliar do Instituto Superior de Economia e Gestão da Universidade Técnica de Lisboa;

- Doutora Inês Maria Galvão Teles Ferreira da Fonseca Pinto, assistente do Instituto Superior de Economia e Gestão da Universidade Técnica de Lisboa.

Dezembro/2011



*"Risk is like fire: If controlled it will help you,
if uncontrolled it will rise up and destroy you."*

Theodore Roosevelt

RESUMO

O objetivo deste trabalho reside na produção de inputs necessários à futura implementação de um modelo de avaliação do risco de crédito de clientes numa empresa comercializadora de energia instalada em Portugal.

A relevância do tema, assume particular interesse no atual contexto, quer porque o processo de liberalização do mercado de comercialização de energia eléctrica se encontra praticamente completo, quer porque a atual crise económica e financeira nacional e internacional tem vindo a abalar a capacidade creditícia das empresas nacionais.

No sentido de aproveitar sinergias, foram testados dois modelos de avaliação do risco de crédito utilizados numa outra empresa do Grupo, que privilegiam a utilização dos rácios económicos e financeiros, como indicador do incumprimento dos clientes.

Estes modelos foram aplicados a uma amostra representativa da carteira real de clientes, constituída de modo equilibrado por clientes “Com Risco” e “Clientes Sem Risco”.

Os resultados obtidos não se revelaram satisfatórios, tendo-se iniciado uma segunda fase do trabalho com a finalidade de definir modelos alternativos fiáveis, capazes de prever o incumprimento de clientes num horizonte temporal de pelo menos um ano.

Na análise dos dados, foi utilizado o Método de Regressão Logística (*Logit*), o qual, a partir de um conjunto de variáveis independentes, estima a probabilidade de ocorrer um certo evento, neste caso, a ocorrência de incumprimento.

O processo de modelação passou por diversas etapas, tendo-se utilizado uma bateria de vinte e oito rácios económicos e financeiros e três outros indicadores financeiros, utilizados em diversos estudos de outros autores, no mesmo âmbito, e testado diferentes “zonas” de segmentação, isto é, várias segmentações da amostra por volume de negócios dos clientes.

No final do processo, obtiveram-se dois modelos para a avaliação do risco de crédito de clientes, que por apresentarem um melhor ajustamento e precisão de entre os vários testes, foram os selecionados.

Os modelos estimados aplicam-se a empresas com volume de negócios anual inferior e superior a um milhão de euros respetivamente e apresentam taxas de acerto que se situam acima dos 60%.

Apesar das taxas de acerto não se revelarem muito elevadas, acredita-se que as mesmas teriam sido substancialmente superiores se a classificação inicial dos clientes da amostra não apresentasse alguns constrangimentos inerentes ao próprio negócio do comércio de energia.

O trabalho culmina com a identificação das atuais limitações e propõe futuros trabalhos complementares e de melhoria.

Palavras Chave: Risco de crédito, Incumprimento, Regressão Logística (*Logit*), Clientes

ABSTRACT

The objective of this work lies in the production of inputs required for future implementation of a model to evaluate the credit risk of customers in a trading power company installed in Portugal.

The relevance of this subject is of particular interest in the current context because the liberalization of electricity market is almost complete and that current economic, financial, domestic and international crisis has shaken the creditworthiness of the national companies.

In order to take advantage of synergies were tested two models of credit risk assessment used in another Group company, that favor the use of economic and financial ratios as an indicator of the failure of customers.

These models were applied to a representative sample of the actual portfolio of clients, consisting in a balanced way by customers “With Risk” and “Without Risk”.

The results had not proved satisfactory, and has started a second phase of work in order to set reliable alternative models, capable of predicting the failure of customers on a time horizon of at least one year.

In the data analysis was used Logistical Method of Regression (Logit), which uses sets of independent variables to estimate the probability of a given event occurring – in this case, failure to comply with payments.

The modeling process went through several stages and it was used a battery of twenty-eight economic and financial ratios. These ratios were used in similar studies by other authors.

Additionally were tested different areas of cut, wich means, different segmentations of the sample by turnover of customers.

At the end of the process were obtained two models for assessing the credit risk of customers which, because they have a better fit and precision, were selected.

The estimated models applies to the companies with annual turnover below and above one million euros and have accuracy rates that are above the 60%.

Despite the accuracy rates do not prove too high, it is believed that they would have been substantially higher if the initial classification of customers of the sample did not present some constraints inherent to the business of energy trading.

The work culminated with the identification of current limitations and propose future studies and further improvements.

Key Words: Credit risk, Failure, Logistical Method of Regression (Logit), Customers

Índice

RESUMO	IV
ABSTRACT	VI
ÍNDICE.....	VIII
LISTA DE FIGURAS	X
LISTA DE QUADROS.....	XI
LISTA DE GRÁFICOS	XIII
AGRADECIMENTOS.....	XIV
INTRODUÇÃO	15
CAPITULO 1 - REVISÃO DE LITERATURA.....	18
1.1. RISCO DE CRÉDITO COMERCIAL	18
1.2. FACTORES DETERMINANTES DO INCUMPRIMENTO.....	18
1.3. METODOLOGIAS DE AVALIAÇÃO DO RISCO DE CRÉDITO	21
1.3.1. <i>Sistemas de “Rating”</i>	21
1.3.2. <i>Sistemas de “Scoring”</i>	22
CAPITULO 2 - DESCRIÇÃO DO MODELO ATUAL	34
CAPITULO 3 - VERIFICAÇÃO DA APLICABILIDADE DO MODELO	41
3.1. CRITÉRIOS PARA CLASSIFICAÇÃO DO RISCO DOS CLIENTES.....	41
3.2. DEFINIÇÃO E CARACTERIZAÇÃO DA AMOSTRA.....	42
3.3. APLICAÇÃO DOS MODELOS À AMOSTRA	48
3.3.1. <i>Resultados Obtidos</i>	48
3.3.2. <i>Conclusão</i>	50
CAPITULO 4 - DEFINIÇÃO DE MODELOS ALTERNATIVOS.....	52
4.1. SEGMENTAÇÃO DA AMOSTRA	52
4.2. SELEÇÃO DAS VARIÁVEIS	55
4.3. METODOLOGIA.....	57

4.3.1. <i>Modelo Logit</i>	57
4.4. RESULTADOS.....	62
4.4.1. <i>Análise Descritiva</i>	62
4.4.2. <i>Modelo de Regressão Logística - Logit</i>	66
CAPITULO 5 - CONCLUSÕES FINAIS	85
5.1. PRINCIPAIS CONCLUSÕES	85
5.2. PROPOSTA DE MELHORIAS.....	89
BIBLIOGRAFIA.....	91
ANEXO 1 – ANÁLISE DISCRIMINANTE.....	95
ANEXO 2 – MODELO Z-SCORE DE ALTMAN	96
ANEXO 3 – ANÁLISE DESCRITIVA DAS VARIÁVEIS	98
ANEXO 4 – MATRIZ DE CORRELAÇÕES DE PEARSON.....	101
ANEXO 5 – MODELOS REJEITADOS	104

Lista de Figuras

Figura 1- Scoring Final – Correção Setorial e Qualitativa	40
Figura 2 – Caracterização População vs Amostra Final.....	47

Lista de Quadros

Quadro 1 - Insolvências - Evolução Anual (Variações Homólogas).....	19
Quadro 2 - Resumo das Técnicas Estatísticas Usadas em Credit Scoring	33
Quadro 3 – Modelo Atual - Rácios Económicos e Financeiros	35
Quadro 4 – Variáveis Críticas	38
Quadro 5 – Escala de Pontuação dos Modelos.....	39
Quadro 6 – Definição dos Segmentos	43
Quadro 7– N.º de Clientes Por Intervalo de Volume de Negócios	46
Quadro 8 – Modelo 1 – Matriz de Classificação	50
Quadro 9 – Modelo 2 – Matriz Classificação	50
Quadro 10 - Tipos de Erro.....	50
Quadro 11 – Variáveis Seleccionadas.....	56
Quadro 12 – Análise Descritiva - Rendibilidade do Ativo	62
Quadro 13 – Análise Descritiva - Rotação Ativo.....	63
Quadro 14 – Análise Descritiva - Aut. Financeira	63
Quadro 15 - Análise Descritiva - Cob. Custos Financeir.	64
Quadro 16 - Análise Descritiva - Liquidez Imediata.....	64
Quadro 17 - Análise Descritiva - Ativos Tangíveis.....	64
Quadro 18 – Medidas de Ajustamento do Modelo.....	68
Quadro 19 – Teste Hosmer and Lemeshow	70
Quadro 20 – Tabela de Classificação ^a	71
Quadro 21 – Amostra de Treino – Matriz de Classificação	74
Quadro 22 – Amostra de Validação – Matriz de Classificação	74
Quadro 23 – Variáveis do Modelo.....	75
Quadro 24 – Probabilidades de Incumprimento	76
Quadro 25 – Medidas de Ajustamento do Modelo.....	77
Quadro 26 – Teste Hosmer and Lemeshow	78

Quadro 27 – Tabela de Classificação ^a	79
Quadro 28 – Amostra de Treino – Matriz de Classificação	82
Quadro 29 – Amostra de Validação – Matriz de Classificação	82
Quadro 30 – Variáveis do Modelo.....	83
Quadro 31 – Probabilidades de Incumprimento	84
Quadro 32 - Matriz de classificação (um ano antes da falência).....	96
Quadro 33 - Matriz de classificação (dois anos antes da falência)	97
Quadro 34 – Análise Descritiva – Clientes “Sem Risco”	98
Quadro 35 - Análise Descritiva – Clientes “Com Risco”	99
Quadro 36 – Matriz de Correlações de Pearson	101
Quadro 37 – Matriz de Classificação – Modelo Sem Segmentação	104
Quadro 38 - Matriz de Classificação – Modelo “Clientes VN < 2 Milhões Euros”	104
Quadro 39 - Matriz de Classificação – Modelo “Clientes VN > 2 Milhões Euros”	104
Quadro 40 – Matriz de Classificação – Modelo “Clientes VN < 10 Milhões Euros”	104
Quadro 41 - Matriz de Classificação – Modelo “Clientes VN > 10 Milhões Euros”	105
Quadro 42 – Matriz de Classificação – Modelo “Clientes VN < 50 Milhões Euros”	105
Quadro 43 - Matriz de Classificação – Modelo “Clientes VN > 50 Milhões Euros”	105

Lista de Gráficos

Gráfico 1 – Status de Clientes.....	43
Gráfico 2 – População Clientes/Segmento	44
Gráfico 3– Amostra Clientes/Segmento.....	44
Gráfico 4 – Amostra – Classificação de Clientes.....	44
Gráfico 5 – Percentagem de Clientes por Localização Geográfica.....	44
Gráfico 6 – Percentagem de Clientes por Atividade Económica.....	45
Gráfico 7 – Amostra – Classificação de Clientes Por Segmento	46
Gráfico 8 – Composição da Amostra Final de Clientes	48
Gráfico 9 – Frequência de Clientes Por Intervalo de VN.....	53
Gráfico 10 – Representatividade dos Segmentos Por Intervalo de VN	54

Agradecimentos

Em primeiro lugar, gostaria de agradecer à empresa que me acolheu há dois anos e meio atrás e me tem permitido crescer pessoal e profissionalmente, tendo contribuído para isso, a possibilidade de realizar o presente trabalho.

Por me terem proporcionado todas as condições para a realização dos estudos, facultando a utilização de dados da empresa, pela transmissão de informação, pelo interesse e confiança demonstrados, o meu muito obrigado.

Agradeço ao Eng. Henrique oliveira pelo conhecimento e disponibilidade sempre demonstrados.

Para todos aqueles que são a minha fonte de inspiração diária, amigos, pais e irmã, pela ausência, pela paciência e pelas palavras de incentivo e amizade, o meu muito obrigado.

Ao meu João, por ter sido o mais afetado neste período conturbado, mas sobretudo pelo seu apoio e paciência incondicionais, agradeço e dedico este trabalho.

Introdução

Com vista a dotar a União Europeia de um setor energético eficiente, competitivo e racional, a Comissão das Comunidades Europeias iniciou em 1990 o processo de liberalização deste setor com a publicação de algumas diretivas relativas à transparência dos preços e ao trânsito de eletricidade e do gás.

O processo de liberalização do setor de comercialização de energia, em Portugal, processou-se de modo gradual, tendo-se iniciado em 2000, com a abertura do mercado aos segmentos de alta tensão, muito alta tensão e média tensão, ficando concluído no ano de 2006, quando foi estendido o espectro de clientes que pode escolher o seu fornecedor de eletricidade aos consumidores residenciais. Porém, o mercado livre de comercialização de energia tem-se mantido concorrente do mercado regulado, e assim continuará até ao final do ano de 2011.

O operador comercial que opera no mercado nacional, vê-se confrontado a reagir ao desenvolvimento do mercado onde se insere. Onde antes existia previsibilidade e certeza, na fase pós regulatória, passa a reinar a turbulência incutida pela concorrência.

A estratégia comercial tende a adaptar-se, seguindo a linha do crescimento orientado aliado a um risco controlado.

O risco pode assumir diversas faces, umas ligadas à parte operacional do negócio, outras, mais relacionadas com a dinâmica do próprio mercado onde a comercializadora atua, como é o caso do risco de crédito comercial ou de clientes.

O risco de contraparte assume um papel preponderante na avaliação dos resultados da empresa, na realidade, a empresa que se pretende líder de mercado deve manter o risco vigiado e controlado.

O risco de crédito de clientes torna-se mais preocupante num quadro económico e financeiro débeis, quer a nível nacional, quer internacional, como o que é vivido atualmente. O setor empresarial vê-se visivelmente afetado com a redução dos apoios do Estado (subsídios, bonificações, deduções fiscais e aumento dos impostos), e dificuldades de financiamento junto do setor bancário, já para não falar no efeito causado pelo desemprego, por meio da diminuição do consumo e da redução na procura.

As questões enunciadas, culminam em problemas de liquidez graves, que levam as unidades produtivas a não conseguir cumprir os seus compromissos de crédito. Incapazes de solver as suas dívidas, as empresas acabam por declarar falência e encerrar definitivamente a sua atividade.

As previsões para o próximo ano, também não apontam para um retrocesso do quadro descrito. Segundo o “Estudo de Insolvências de Empresas – Portugal – 1º Trimestre 2011/2010” da Coface, o número de ações/decisões de insolvência registados nos meses de Janeiro e Março 2011, cresceram 9,7% face a igual período de 2010. As previsões da OCDE revelam uma continuidade do quadro atual, apontando para uma recessão de 2,1% em Portugal para este ano e de 1,5% no próximo. Ainda nas previsões de Primavera, a OCDE prevê que em 2012, Portugal atinja uma taxa de desemprego que deverá rondar os 13% e a dívida pública deverá atingir os 115% do PIB, bem acima de toda a riqueza produzida no país ao longo de um ano.

Os custos para a empresa, associados à inclusão de “maus” clientes na sua carteira são bastante elevados. Conforme refere Gamelas (2005), os credores obtêm apenas 10% das dívidas a que tinham direito, esperando em média 10 anos para receberem os créditos, e só os custos com a liquidação absorvem, em média, 29% do valor de realização dos ativos em liquidação e que as dívidas reclamadas, em média, são superiores em 37% ao passivo declarado em balanço contabilístico.

Os resultados obtidos num estudo elaborado recentemente sobre a exposição atual da empresa ao risco de incumprimento dos clientes, revelou que no período compreendido entre Dezembro 2010 e Agosto 2011, a dívida vencida de clientes sofreu um ligeiro aumento, na ordem dos 4%. De realçar, no entanto, que a dívida vencida representa 24% do total da dívida de clientes, concentrando-se sobretudo nas grandes maturidades, ou seja, 63% da dívida vencida está concentrada nas maturidades acima dos 121 dias, o que poderá ser apontado como um aspecto preocupante, uma vez que se considera que a probabilidade de recuperação da dívida, será tanto menor, quanto maior a sua maturidade. Outras conclusões do referido trabalho, apontam para um acréscimo (8,17%) do número de clientes em contencioso desde Dezembro 2010, sendo que 56,9% dos casos de contencioso se devem a processos de insolvência¹.

¹ Dada a confidencialidade da informação constante no estudo, o mesmo não constará na lista de referências bibliográficas

Perante a importância demonstrada do tema para a empresa, o presente trabalho projeto, pretende oferecer diversos contributos para a futura implementação de um modelo de avaliação do risco de crédito na comercializadora de energia que opera no mercado nacional.

Mais concretamente, os objetivos a alcançar são os seguintes:

1. Levantamento de alguns requisitos essenciais, nomeadamente a caracterização e representatividade dos clientes na carteira, por dimensão, atividade, localização geográfica e segmento de negócio.

Esta primeira parte do trabalho será realizada em simultâneo com a seleção da amostra de clientes, que será utilizada na fase seguinte.

2. Verificação da aplicabilidade de dois modelos de risco de crédito utilizados numa outra empresa do Grupo, com base numa amostra de clientes representativa da carteira real.

Os modelos serão aplicados a uma amostra de clientes da empresa que participa no mercado nacional, posteriormente, com base nas classificações de risco de crédito obtidas para cada empresa da amostra, será efetuado um exercício de comparabilidade entre os resultados da aplicação do modelo e a situação atual das empresas em carteira, permitindo aferir sobre a aplicabilidade do modelo.

Os modelos serão testados somente na sua vertente quantitativa, uma vez que esta é a sua principal componente, e não se dispõe de informação suficiente para testar as restantes variáveis de cariz qualitativo.

3. Estimar um ou vários modelos quantitativos de avaliação do risco de crédito de clientes, utilizando rácios económicos e financeiros e recorrendo a uma metodologia estatística eficiente e de fácil interpretação.
4. Concluir sobre a razoabilidade e adequabilidade dos modelos estimados face aos testados, evidenciando eventuais limitações e propondo futuros estudos e melhorias.

Capítulo 1 - Revisão de Literatura

1.1. Risco de Crédito Comercial

Segundo Bessis (1998), o risco de crédito é definido pelas perdas geradas por um evento de default do devedor, ou pela deterioração da sua qualidade de crédito. O autor cita como exemplos de situações que podem caracterizar um evento de default do devedor, as seguintes: Atraso no pagamento de uma obrigação, o não cumprimento de uma cláusula contratual restritiva, o início de um procedimento legal como a insolvência/falência, ou a inadimplência de natureza económica, que ocorre quando o valor económico dos ativos da empresa se reduz a um nível inferior ao das suas dívidas, indicando que os fluxos de caixa esperados não são suficientes para liquidar as obrigações assumidas.

O risco de crédito é algo que está presente no quotidiano de qualquer empresa, seja uma empresa da área financeira, uma empresa de serviços, comercial ou industrial.

De acordo com Silva (2004), em termos gerais, há três tipos de crédito, o que é concedido por empresas comerciais, nas vendas ao público e a empresas, e o que é concedido por instituições financeiras a particulares e a empresas.

Brealey & Myers (1998) emprestam ao risco de crédito comercial uma visão economicista, denunciando a crescente representatividade deste tipo de crédito no balanço das empresas. Os autores evidenciam que as empresas que não efetuam o pagamento a prazo das suas aquisições estão, efetivamente a contrair um empréstimo junto dos respetivos fornecedores. Estas “dívidas” refletem-se nos balanços das empresas compradoras como contas a pagar, constituindo esta rubrica a origem mais importante do financiamento de curto prazo, representando quase o dobro dos empréstimos de curto prazo contraídos junto de bancos ou de outras instituições.

1.2. Factores Determinantes do Incumprimento

As estatísticas nacionais e internacionais têm revelado um aumento das taxas de incumprimento nos últimos anos: Segundo Jonathan N. Crook et al. (2007), entre 1970 e

2005 o volume de crédito ao consumo por saldar nos E.U.A cresceu 231% e o volume dos empréstimos bancários securitizados cresceu 705%. Por seu lado, em Portugal, “O crédito malparado das sociedades não financeiras atingiu os 6.162 mil milhões de euros em Novembro de 2010, tendo subido cerca de 16% em termos homólogos, e representa já 5,2% do total.” (Diário Económico, 2011).

De um modo geral, a literatura relacionada com a temática do risco de crédito comercial, aponta a insolvência como a principal causa de incumprimento das empresas, relativamente aos compromissos que assumem junto dos credores.

O Decreto-Lei n.º 53/2004, de 18 de Março, que aprova o Código da Insolvência e da Recuperação de Empresas, estabeleceu que é “considerado em situação de insolvência o devedor que se encontre impossibilitado de cumprir as suas obrigações vencidas”. Segundo o mesmo articulado, “a insolvência não se confunde com a falência, tal como atualmente é entendida, dado que a impossibilidade de cumprir obrigações vencidas, em que a primeira noção fundamentalmente consiste, não implica a inviabilidade económica da empresa ou a irrecuperabilidade financeira postuladas pela segunda”.

Segundo Pires Manso (2010), os dados demonstram que o número de empresas insolventes em Portugal, tem vindo a crescer incessantemente desde há alguns anos, tendo-se registado um crescimento de 7,75% no 1º trimestre de 2010 face ao período homólogo de 2009.

Quadro 1 - Insolvências - Evolução Anual (Variações Homólogas)

Anos	2008	2009	2010
N.º Insolvências	821	1122	1209

Fonte: Manso, José, 2010, “Insolvências de Empresas”, Observatório Para o Desenvolvimento Económico e Social

O estudo avança ainda, que as principais causas desta pronunciada evolução devem-se sobretudo aos seguintes aspectos:

- Crise financeira internacional que contribuiu para o agravamento da crise económica nacional;
- Crise do subprime ou imobiliária;

- O fenómeno da globalização: A abertura das fronteiras ao comércio de economias mais competitivas, levou ao encerramento de um importante número de empresas pertencentes a setores tradicionais;
- A concorrência das grandes superfícies que tem vindo a provocar o encerramento de inúmeras empresas do comércio por grosso e sobretudo a retalho (caso do comércio tradicional);
- Encerramento de muitas PME's e empresas familiares dos setores agrícola e pecuário devido à aplicação de uma política agrícola comum da UE limitadora do potencial produtivo destes setores.

Outros autores como Nogueira Leite (2009) também identificam o problema da competitividade como responsável pelo aumento de insolvências na indústria, que se agrava com a quebra da atividade económica sentida em Portugal e no estrangeiro.

Para o economista e professor da Universidade do Porto, João Loureiro, citado num artigo de jornal, "(...)a tendência do crédito malparado foi visível durante todo o ano de 2010 e promete não parar tão cedo, o que se deve sobretudo à situação da economia em que se vai viver, e porque existem ainda muitas incertezas sobre a evolução da economia." (Loureiro, J. 2011).

Para Altuzarra C. (2011), economista da Coface, a crise chegou a Portugal e o Governo foi obrigado a implementar medidas de austeridade, que se intensificarão por via do acordo estabelecido com a "troika" e a UE, prevendo-se um quadro de recessão económica para o ano de 2011, que se espera na ordem dos 1,3%, e que se deverá arrastar para o próximo ano. O problema reside efetivamente, no facto das empresas necessitarem de crédito, para a economista da Coface este problema não se coloca para as empresas que dispõem de fundo de maneo, no entanto, a maioria das empresas portuguesas necessita de crédito. Outro problema apontado pela economista, é o facto da economia portuguesa estar muito centrada na UE, sendo necessário desenvolver outros mercados.

Para Gabás (1990) as causas mais importantes que podem contribuir para um estado de insolvência nas empresas são:

1. Causas com origem externa:

- De mercado: Concorrência excessiva e forte queda da procura;
- Da envolvente política, económico e social: Fase depressiva do ciclo económico, situações de crise, política económica do governo, mudanças sociais radicais e significativas.

2. Causas com origem interna:

Ineficácia da direção, estratégias erradas ou inadequadas, sistema produtivo ineficiente, investimentos improdutivos, excessivo endividamento agravado em certas épocas por elevadas taxas de juro, final do ciclo de vida do produto, fracasso de empresas do mesmo grupo; problemas internos não resolvidos; alta morosidade.

3. Causas especiais: Novas empresas apresentam uma taxa de “mortalidade” muito elevada nos primeiros anos de atividade.

Para Richardson, Kane & Lobingier (1998), também as condições macroeconómicas influenciam o aparecimento de situações de crise para a empresa, podendo levar a situações de falência.

Bonfim (2006) conclui, que embora as probabilidades de incumprimento sejam determinadas essencialmente pelas características específicas de cada empresa, existe igualmente uma relação entre a evolução da atividade macroeconómica e as taxas de incumprimento.

1.3. Metodologias de Avaliação do Risco de Crédito

Carvalho das Neves (2005) classifica os sistemas clássicos de análise de risco de crédito em sistemas de “scoring” (pontuação) e sistemas de “rating” (notação de crédito). Mais recentemente têm-se desenvolvido outros modelos de análise de risco de crédito, tendo por base, nomeadamente:

- a) Modelos que utilizam a inteligência artificial, como sejam os expert systems e as redes neuronais;
- b) Modelos que utilizam informação de mercado, como seja a estrutura temporal da taxa de juro e as taxas de mortalidade e migração do crédito;
- c) Modelos que utilizam a teoria das opções na avaliação do risco de incumprimento.

1.3.1. Sistemas de “Rating”

Segundo Falkenstein (2000), os sistemas de “rating” visam classificar uma emissão obrigacionista ou outros títulos de dívida numa escala de notação de risco e refletem um

juízo de valor sobre a capacidade de pagamento atempada do serviço da dívida da operação em análise. A classificação final depende da percepção que o analista tenha sobre a situação global da empresa e da sua capacidade de cumprimento da dívida em causa.

No “rating” a informação qualitativa ou quantitativa é tratada de forma a produzir no analista a percepção de qual o grupo de risco a que a empresa em causa pertence, a partir de uma definição prévia de tipo de grupo de risco. Por exemplo, a Standard & Poor’s utiliza 10 graus diferentes, que podem variar entre “Aaa” e “C”, para classificar o crédito do melhor para o pior, respetivamente.

1.3.2. Sistemas de “Scoring”

Os sistemas de “scoring” ou de pontuação do risco de crédito, “(...)são sistemas que utilizam rácios e outros indicadores para classificar, de forma automatizada, as empresas por grau de risco e baseiam-se, normalmente, em informações históricas” (Carvalho das Neves, 2005, p. 202). Estes modelos podem ser teóricos ou empíricos com desenvolvimentos mais ou menos sofisticados do ponto de vista técnico.

L. C. Thomas (2009) refere que estamos na era da terceira geração dos modelos de credit scoring, denominados por profit scoring, onde se pretende avaliar não só o perfil de risco do solicitante de crédito, mas igualmente, a probabilidade do candidato ao crédito dar lucro à instituição, não sendo a avaliação apenas baseada no risco.

Para este autor o principal objetivo do scoring é melhorar o processo de seleção de bons clientes de modo a reduzir perdas futuras. Dado o seu sucesso, os sistemas de scoring tornaram-se um factor de decisão chave e uma ferramenta de suporte à decisão, imprescindível na quantificação e gestão do risco.

Segundo Chorão (2005), os métodos utilizados em credit scoring incluem várias técnicas estatísticas e de investigação operacional, sendo as mais utilizadas a regressão logística, a análise discriminante e as árvores de decisão.

Para Carvalho das Neves (2005) as técnicas estatísticas podem ser paramétricas ou não paramétricas. Das técnicas paramétricas destacam-se as análises univariadas e multivariadas. Na análise multivariada as mais utilizadas na análise do risco de crédito são a análise discriminante, o “logit” e o “probit”.

1.3.2.1. Análise Univariada

Para Cook & Nelson (1998), esta metodologia constitui um processo estatístico que permite a análise e identificação das variáveis que individualmente possuem um maior poder explicativo face à possibilidade de incumprimento.

As técnicas estatísticas da Análise Univariada têm como objetivo a descrição de cada uma das variáveis, recorrendo a frequências, medidas de síntese (estatística descritiva) e a representações gráficas dos dados.

Num dos primeiros estudos levados a cabo sobre a insolvência baseada na estrutura financeira da empresa, Beaver (1966) analisa a tendência de sete rácios financeiros (cash flow/Passivo; Resultado Líquido/Ativo; Passivo/Ativo; Fundo de Maneio/Ativo; Liquidez Geral; Intervalo de Segurança; Ativos), por serem os mais utilizados pelos analistas financeiros naquela altura, durante 5 anos antes da falência. O autor compara os valores dos rácios obtidos para uma amostra de 79 empresas insolventes, em relação a outras 79 empresas em situação considerada normal, tendo estas últimas sido selecionadas fazendo o “matching” por indústria, região e dimensão das empresas insolventes da amostra.

Beaver (1966) chega a resultados bastante encorajadores com apenas 9% de empresas classificadas incorretamente.

Estes resultados embora interessantes têm uma série de problemas que convém realçar:

a) Os erros de previsão tipo I e tipo II não têm o mesmo custo para o decisor (Altman et al., 1977). Para a generalidade dos credores é mais oneroso ocorrer a falência quando estes a não previam (Erro tipo I), que é precisamente onde a previsão de Beaver mais falha;

b) A amostra de Beaver subdivide-se em partes iguais de empresas falidas e não falidas. No universo o número de empresas não falidas é maior do que o número de empresas falidas e o erro de previsão tende a ser maior na prática do que o estimado pelo modelo;

c) A seleção dos rácios não se baseou em qualquer teoria da falência, mas por serem bastante utilizados na prática. Um dos inconvenientes é que, por serem muito utilizados por analistas, são também aqueles que os processos de contabilidade criativa procuram “mascarar” (Zavgren, 1983);

d) Na Análise Univariada assume-se, segundo Cook & Nelson (1998) que “*a single variable can be used for predictive purposes*”, porém, o teste utiliza separadamente os rácios quando, na produção dos relatórios de análise financeira o analista procura interligá-

los para compreender a evolução e a situação financeira da empresa. De facto, um rácio apenas, num determinado momento, diz muito pouco ao analista financeiro.

O modelo univariado proposto por Beaver atingiu, segundo Sheppard (1994), “*a moderate level of predictive accuracy*”. A análise univariada permitiu identificar factores relacionados com a falência, no entanto, segundo Stickney (1996), aquela análise não proporcionou uma relevante medida do risco.

1.3.2.2. Análise Multivariada

A análise multivariada, também conhecida como análise discriminante múltipla tentou, conforme referem Cook & Nelson (1998), ultrapassar os problemas resultantes da utilização de variáveis isoladamente.

De acordo com Hair et al. (1998), a Análise Discriminante e a Regressão Logística, enquadram-se na classe de métodos estatísticos multivariados de dependência, uma vez que relacionam um conjunto de variáveis independentes com uma variável depende categórica.

1.3.2.2.1 – Análise Discriminante Linear

A análise discriminante linear, apesar das insuficiências técnicas referidas em vários estudos de aplicação às finanças, como seja o pressuposto de que as variáveis utilizadas para caracterizar os grupos têm distribuição normal, que as matrizes de dispersão (variância-covariância) são iguais para todos os grupos, ou o pressuposto de que os grupos são discretos e identificáveis, é a mais utilizada neste tipo de trabalhos de investigação. Prova disso, é a sua já longa utilização pelos principais bancos centrais europeus, na avaliação do risco de crédito dos seus clientes, conforme refere Wende Waltraud, membro do departamento de crédito do Banco Central Alemão, na conferência internacional “European Committee of Central Balance Sheet Data Offices”, realizada em Paris em 1997: O Deutsche Bundesbank desde os anos setenta, o Banco Central Francês e Italiano, desde a década de oitenta e o Banco de Inglaterra desde 1993. As principais vantagens apontadas para a sua utilização são a eficiência, simplicidade e fácil interpretação no processo de avaliação do crédito, para além de revelar uma técnica estatisticamente muito estável.

Também de referir, os trabalhos de Altman nos EUA enquanto precursor na investigação do risco de falência e, posteriormente, pela comercialização deste tipo de informação, conforme será referenciado seguidamente.

Altman, considerado o pioneiro na aplicação da Análise Discriminante (Anexo 1) ao problema da Previsão da Falência, propôs em 1968, um modelo baseado na análise discriminante que viria a ser conhecido como Z-score.

O modelo de Altman (Anexo 2) caracteriza-se pela combinação linear de cinco rácios, medidas de rentabilidade e risco, constituindo um previsor capaz de discriminar entre empresas falidas e não falidas com uma elevada percentagem de êxito nos dois anos anteriores à falência.

Altman (1968) selecionou um conjunto de rácios – característica quantitativa – e ponderações atribuídas de modo a produzir uma classificação Z, isto é, um indicador de síntese (*score*) cujos valores permitissem diferenciar, da melhor forma possível, a população de empresas “normais”, da população de empresas “de risco” – característica qualitativa.

A função apresentada em 1968 e posteriormente modificada, devido a diversos contributos, foi elaborada tomando por suporte uma amostra de 66 empresas, a qual inclui 33 em processo de falência, no período entre 1946 e 1965, e 33 em situação “normal”, e um conjunto de 22 rácios.

Após o tratamento estatístico dos dados, Altman (1968) considerou que existiam 5 rácios, cujos valores diferiam significativamente entre empresas falidas e não falidas.

Além de permitir classificar as empresas por grupos de maior ou menor risco, este modelo permite ainda, identificar qual o rácio que melhor discrimina os grupos.

Este modelo provou ser capaz, para a amostra de empresas, de prever com uma precisão de aproximadamente 94% a falência no espaço de um ano e com uma precisão de mais de 70% a falência no espaço de dois anos.

Este modelo levanta, no entanto, um problema de aplicabilidade, uma vez que exige o conhecimento do valor de mercado da empresa, valor que não é facilmente obtido nos casos em que a empresa não está cotada na bolsa. No sentido de ultrapassar este problema, Altman defende que o modelo deverá ser novamente estimado, utilizando o valor dos capitais próprios em vez do valor de mercado.

Apesar dos resultados positivos do estudo de Altman (1968), o seu modelo tinha uma “debilidade”, assumia as variáveis na amostra como tendo distribuição normal. Segundo Sheppard (1994), se as variáveis não têm distribuição normal, o método utilizado pode resultar na seleção de um conjunto não apropriado de previsores.

De referir ainda, que os rácios escolhidos por Altman não assentam em nenhuma teoria, mas na eficiência estatística do modelo, pelo que a amostra utilizada influencia aquela escolha.

O trabalho inicial de Altman foi revisto e melhorado por muitos investigadores. Os principais trabalhos a referir, foram os de Deakin (1972), Blum (1974), Edmister (1972), Libby (1975), Scott (1981) e Taffler (1982).

Também Altman desenvolveu um modelo Z-score revisto (com coeficientes revistos) que abandonou algumas variáveis anteriormente utilizadas, substituindo-as por uma nova variável que passou a ser “capitais próprios sobre passivo total” (net worth (book value)/ total liabilities).

Por volta de 1977, Altman, Haldman e Narayanan desenvolveram, juntamente com uma empresa financeira privada (Zeta Services, Inc.) um outro modelo revisto, com sete variáveis, baseado na amostra combinada de 113 fabricantes e retalhistas. O modelo Zeta é supostamente mais preciso na classificação de empresas falidas entre os 2 e 5 anos anteriores. O modelo Zeta baseia-se nas seguintes variáveis: Rendibilidade do ativo (Resultado Operacional sobre o Ativo); Estabilidade da rendibilidade (medida pelo erro standard normalizado em torno da tendência de 10 anos da rendibilidade do ativo, esta métrica funciona como um indicador do risco de negócio); Serviço da dívida (rácio de cobertura dos custos financeiros - resultado operacional sobre custos financeiros); Rendibilidade acumulada – medida como “resultados transitados e reservas sobre activo”, segundo Altman *et al.* (1977) um dos indicadores mais relevantes na discriminação das empresas; Liquidez – medida que permite detectar possíveis problemas de tesouraria; Capitalização – medido pelo rácio de Autonomia Financeira com base na média dos valores de mercado de 5 anos; Dimensão – medida como o logaritmo do total dos ativos tangíveis.

1.3.2.2.2 – Modelo Logit/Probit

Para Carvalho das Neves (2005), enquanto a análise discriminante procura, dentro de um conjunto de características, a melhor forma de discriminar os grupos de risco (normalmente dois ou três), os modelos “logit” e “probit”, que se baseiam, respetivamente, numa função logística e numa função normal, procedem de forma “dual”, ao abordarem esta problemática de uma forma semelhante ao modelo de regressão linear, isto é, ao partirem da definição das subpopulações para obterem a função que permite efetuar a discriminação.

É possível identificar mais algum tipo de contingências nos modelos de análise discriminante face à regressão logística, nomeadamente o facto dos resultados da análise discriminante só se considerarem válidos se as variáveis independentes forem normalmente distribuídas, para além disso, ao contrário da análise discriminante, a regressão logística exige um número menor de pressupostos e que são menos rígidos (SPSS, 2003).

Para além destes aspectos, o resultado obtido na análise discriminante corresponde a uma pontuação, não traduzindo nenhum carácter probabilístico, o que leva à preferência pela utilização do Modelo Logístico (Ohlson, 1980).

Devido à utilização da Análise Logit, o modelo é considerado, segundo Lo (1986), “more robust” que a Análise Discriminante, já que é aplicável a outras distribuições que não a normal.

Segundo Ferrando & Blanco (1998), uma das vantagens do modelo “logit”, prende-se com o facto de admitir que as variáveis independentes possam ser categóricas, permitindo que as variáveis explicativas não se reduzam unicamente a rácios económicos e financeiros ou variáveis métricas, possibilitando desta forma a utilização de informação não financeira ou qualitativa.

Hair (1998), também refere algumas razões pelas quais a regressão logística representa uma alternativa à análise discriminante, nomeadamente quando a variável dependente tem apenas duas categorias. Considera que a regressão logística é menos afetada pelas desigualdades variância/covariância e analisa facilmente variáveis categóricas independentes enquanto que na análise discriminante a utilização de variáveis dummy gera problemas com as igualdades variância/covariância.

Outra vantagem da regressão logística é a sua análise probabilística, uma vez que a regressão estima a probabilidade de ocorrer um certo evento, a partir de um conjunto de variáveis independentes ou explicativas.

O modelo “logit” possui também alguns inconvenientes, Mora (1996), refere algumas limitações associadas à análise discriminante, nomeadamente a não consideração das probabilidades prévias e dos erros de classificação, sempre que se utilize como ponto de corte 0,5.

Outros métodos usados para estimar modelos com variáveis dicotômicas são o Probit e o Gompit, são similares e utilizam o método de estimação da máxima verosimilhança

A transformação mais utilizada é a Logit, uma vez que é mais fácil de tratar do ponto de vista computacional.

Em comparação com o modelo probit, o logit tem representação e tratamento matemático mais simples, justificando a sua maior utilização prática.

Para Borooah (2002), à semelhança do que sucede com o logit, o modelo probit apresenta bons resultados quando a variável dependente é binária ou dicotómica. Uma vez que tem por base uma curva normal padrão, o probit não é recomendado quando existirem muitos casos assimétricos ou com outra distribuição que não a normal (Pampel, 2000).

1.3.2.3 Modelos Não Paramétricos

Mais recentemente, tem-se estudado a eficiência de técnicas não paramétricas na classificação das empresas por grau de risco, nomeadamente, através de métodos da vizinhança mais próxima como o de Fix-Hodges, ou o "algoritmo de partição recursiva" de Loftsgaarden-Quesenberry.

A investigação empírica evidencia que os modelos não paramétricos apresentam uma maior eficiência na classificação de empresas em grupos de risco por comparação com a análise discriminante, nomeadamente o modelo de Fix-Hodges (Stein & Ziegler, 1984).

Numa abordagem mais geral onde se compara a análise discriminante linear e quadrática, o "logit", a programação matemática, a partição recursiva e um modelo teórico com atributos múltiplos de uma instituição bancária americana, Srinivasan & Kim (1987) tiram as seguintes conclusões: Os modelos não paramétricos são, na generalidade, desconhecidos no mundo empresarial, quer porque estes algoritmos não integram os "packages" de maior divulgação quer também porque os estudos comparativos que se referiram evidenciam, em termos metodológicos, algumas lacunas significativas. A primeira limitação surge no facto de apenas se tratarem casos de pequenas amostras particulares (e portanto de difícil generalização), para além disso, nalguns casos a definição das subpopulações não é feita com base num critério objetivo.

1.3.2.4 Modelos de Inteligência Artificial

Nos anos 90 começaram a surgir metodologias baseadas na aplicação de redes neuronais, um dos domínios em que a inteligência artificial se tem mostrado mais fecunda em produzir modelos teoricamente aplicáveis.

Segundo Carvalho das Neves et al. (1998), a referência mais completa às aplicações das redes neuronais a problemas de finanças que se encontrou foi o excelente livro de “readings” editado por Trippi & Turban (1993), no entanto, é possível encontrar alguns trabalhos mais recentes como por exemplo Altman et al. (1994) ou Bardos & Zhu (1997).

Com base nestas referências bibliográficas, (Carvalho das Neves et al., 1998, p.13) definem uma rede neuronal como “(...)uma tecnologia de processamento da informação que se inspira no funcionamento do cérebro humano, isto é, que se baseia numa rede de unidades base, designadas por neurónios, que reagem a estímulos, produzindo assim um determinado “output”. Cada um destes neurónios recebe como estímulo determinada informação primária e/ou “output” originado por outros elementos da rede.”.

Os autores que desenvolvem as redes neuronais costumam apresentar vantagens destas metodologias em termos do número de casos bem classificados. Bardos & Zhu (1997) ilustram esta concepção, trabalhando com uma amostra com mais de 2000 observações e obtendo melhores resultados com as redes neuronais, mas como os próprios implicitamente reconhecem, ainda se está mais numa fase de investigação do que de aplicação prática.

Outros analistas como Altman et al. (1994) partem mais duma visão económica e, embora reconheçam as potencialidades destas novas ferramentas, não deixam de as considerar em termos complementares, assinalando a dificuldade na validação do modelo obtido, à luz dos diferentes rácios introduzidos, a grande quantidade de informação necessária, nem sempre disponível para o analista, e o facto de não haver uma percepção de como o sistema poderá reagir a alterações estruturais. Naturalmente que esta última crítica também é extensível às metodologias “tradicionais”, mas o facto de se conseguir interpretar economicamente as estimativas obtidas para os parâmetros permitirá sempre alguma flexibilidade a este nível.

Klimasaukas (1993) aponta os seguintes problemas: Existe uma dependência acrescida das redes neuronais em relação à dimensão das amostras disponíveis, cerca de 30 a 40 mil observações para a fase de aprendizagem, a fase de transformação da informação é algo heterogénea, uma vez que integra, não só a transformação de variáveis qualitativas em variáveis “dummies”, como também, uma “calibração” das variáveis quantitativas que devem ser postas numa escala de 0 a 1, nomeadamente, através do recurso à logística. A

fase de seleccionar e testar uma arquitectura para a rede é apontado como o ponto mais controverso das metodologias baseadas em redes neuronais, por um lado, esta fase apresenta dificuldades práticas de implementação quer em termos do “software” quer do tempo necessário para a aprendizagem.

Também Carvalho das Neves et al.,1998, p.16, partilham a mesma visão das limitações das redes neuronais, dizendo que “(...) é nossa opinião que se está perante uma metodologia nova e promissora mas em termos complementares às ferramentas estatísticas existentes que, já por si, deverão sempre ser sujeitas ao crivo de uma análise individualizada para os casos mais sensíveis.”.

1.3.2.5 Modelos de Mercado

1.3.2.5.1 – Spread da Taxa de Juro

Este modelo utiliza os prémios de risco inerentes às taxas de rendibilidade até à maturidade (yield) das obrigações das empresas para avaliar o risco de crédito e a probabilidade de incumprimento (Carvalho das Neves, 2005). As agências de rating classificam as empresas por grupos de risco. AAA, AA, A e BBB são créditos considerados de qualidade, BB, B e CCC são categorias de crédito já consideradas em risco e CC, C e D são considerados créditos especulativos e, portanto, de alto risco. O princípio desta análise passa por construir curvas de rendibilidade até à maturidade para as obrigações do tesouro e para as obrigações das empresas classificadas nestes grupos de rating. A diferença (spread) entre as taxas de rendimento (yield) das obrigações em causa e das obrigações do tesouro evidenciam o prémio de risco que o mercado atribui a essas obrigações.

Segundo Fons (1994), este tipo de análise pode ser facilmente desenvolvida para situações mais realistas, nomeadamente, em que exista mais do que um período e o credor espera perder apenas parcialmente o crédito e não a totalidade.

1.3.2.5.2 – Taxas de Mortalidade e Migração do Crédito

Conforme explica Carvalho das Neves (2005), neste método, em vez de estimar as taxas de incumprimento a partir da estrutura temporal das taxas de juro, o analista vai ter em conta os dados históricos de incumprimento e as taxas de mortalidade do crédito de idêntica qualidade.

A taxa de mortalidade marginal no primeiro ano é a probabilidade de incumprimento do devedor durante o primeiro ano do período de reembolso. A taxa de mortalidade marginal

no segundo ano é a probabilidade de incumprimento do devedor neste ano, dado que cumpriu as suas obrigações financeiras no primeiro ano. Assim, para uma obrigação classificada pela agência de rating como BBB tem-se:

$$TMM_i = \frac{VI_{(BBB)i}}{V_{(BBB)i}}$$

Em que:

TMM_i – Taxa de mortalidade marginal no ano i ;

$VI_{(BBB)i}$ – Valor das obrigações BBB que entraram em incumprimento no ano i de vida;

$V_{(BBB)i}$ – Valor das obrigações BBB em cumprimento no início do ano i .

A partir desta informação podem calcular-se as taxas de mortalidade acumulada (TMA) num determinado período (1,2,3,...,n anos) subtraindo à unidade o produto das taxas de sobrevivência da população:

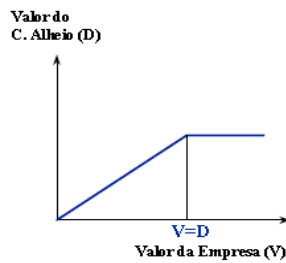
$$TMA_i = 1 - \prod_{i=1}^n (1 - TMM_i)$$

Para além disso é utilizado o conceito de migração que evidencia a mudança de classificação de rating das obrigações de ano para ano. Atualmente a Standard & Poors e a Moody's produzem matrizes de transição que evidenciam a probabilidade de migração do crédito entre diferentes ratings, com este tipo de informação é possível estimar o risco de crédito.

A dificuldade de aplicação destas metodologias de mercado, ao tecido empresarial português, resulta do número limitado de emissões obrigacionistas e do número de empréstimos obrigacionistas com atribuição de rating.

1.3.2.6 Modelo das opções na avaliação do risco de crédito

Na sequência dos trabalhos pioneiros de Black & Scholes (1973) e Merton (1974), o crédito da empresa passa a ser visto como uma “put” sobre os ativos da empresa. Os credores detêm uma opção (put) sobre os ativos que a podem exercer em caso de incumprimento por parte do devedor. No caso de o valor dos ativos ser superior ao valor da dívida, a administração da empresa pagará normalmente o capital e os juros.



Fonte: Adaptado de Carvalho das Neves & Andrade e Silva, 1998

Assim o valor do crédito será $D = \text{MAX}(D, 0)$.

Mas no caso de o valor dos ativos cair abaixo do valor do passivo a empresa entra em insolvência e os credores têm o direito de pôr a empresa em falência recuperando o crédito através da liquidação dos ativos da empresa. Na situação extrema em que os ativos não têm valor, os credores perderão integralmente o valor do capital em dívida e respetivos juros.

Merton (1974) demonstrou que o credor deveria ajustar o prémio de risco em função das variações do endividamento(d) e do risco inerente ao valor dos ativos (σ^2). Quanto maior o endividamento e o risco dos ativos de um devedor, maior o risco de crédito.

De acordo com Saunders (1997) e Boudoukh & Saunders (1998) muitos dos bancos americanos de maior dimensão utilizam este modelo para analisar as frequências das expetativas de risco de incumprimento (EDF - expected default risk frequency). A EDF está concebida para converter a informação contida na evolução dos preços de cotação em medidas de risco de incumprimento, pressupondo uma relação direta entre as seguintes variáveis:

- a) O valor de mercado das ações e o valor de mercado do ativo da empresa devedora;
- b) O risco e valor dos ativos e o risco de incumprimento da sociedade devedora.

Em termos gerais o que o modelo faz é determinar o grau de cobertura da dívida pelos ativos a valor de mercado, tendo em consideração a volatilidade no valor dos mesmos, determinando a probabilidade de o ativo cair abaixo do valor da dívida.

“Para além dos problemas que o modelo tem quanto aos pressupostos, acresce ainda em Portugal, a dificuldade de implementação pelo facto do número reduzido de empresas cotadas” (Carvalho das Neves et al., 1998, p.21).

A tabela que se segue apresenta um resumo dos métodos utilizados:

Quadro 2 - Resumo das Técnicas Estatísticas Usadas em Credit Scoring

Modelo	Principais Técnicas	Resumo
Regressão Linear	Minimos quadrados ordinários	Adequada em situações em que a variável resposta é contínua
Análise discriminante	Distância de Mahalanobis	Classifica os objectos em grupos pré-definidos, minimizando a variância
Regressão Logística	Estimadores de máxima verossimilhança	Adequada em situações em que a variável resposta é binária
Árvores de decisão	Chaid	Utiliza a estrutura da árvore para maximizar a variância entre-grupos
Redes Neurais	Perceptrão multicamada	Técnica de inteligência artificial. Os resultados são difíceis de explicar.
Programação Linear	Método simplex	Muito utilizado na optimização de alocação de recursos

Fonte: Adaptado de Raymond (2007)

Segundo a análise de Raymond (2007), devido à sua flexibilidade e facilidade de utilização, os modelos de probabilidade linear continuam a ser muito utilizados. Estes modelos são muito utilizados em instituições onde o credit scoring tem uma longa história, ou onde a metodologia existente está bem enraizada. Pelo contrário, a regressão logística domina nas instituições onde o credit scoring foi introduzido mais tarde, quer devido às propriedades estatísticas conhecidas quer pela maior transparência e interpretabilidade que introduz no processo de decisão. Os modelos de scoring baseados na regressão logística permitem facilmente identificar estatisticamente as variáveis que mais contribuem para a rejeição do cliente.

As técnicas não paramétricas, têm sido igualmente utilizadas em credit scoring, com algum sucesso. Destacam-se as árvores de decisão e métodos de inteligência artificial, como as redes neuronais, algoritmos genéticos, e método do vizinho mais próximo.

Para Chorão (2005) qualquer que seja a técnica utilizada é correto afirmar que a indústria financeira pretende em primeiro lugar modelos que tenham interpretabilidade e transparência e, em segundo lugar modelos que sejam facilmente implementáveis.

Capítulo 2 - Descrição do Modelo Atual

O modelo de análise do risco de crédito de clientes empresariais, atualmente utilizado numa outra empresa do grupo, instalada num espaço geográfico estrangeiro, tem-se revelado eficiente no propósito para o qual foi criado, apresentando taxas de sucesso elevadas. O seu processo de desenvolvimento baseou-se num processo de construção de um “expert system”, aliado à utilização de algumas técnicas de tratamento estatístico de dados e ao conhecimento técnico e empírico dos seus autores. Trata-se de uma metodologia atual (implementada recentemente), que carece de relativamente poucos inputs e cujos resultados são de fácil interpretação.

Apesar das semelhanças entre as duas empresas, nomeadamente no que toca aos bens comercializados e ao tamanho da carteira de clientes, existem diferenças como sejam, a localização geográfica, a diversidade do tipo e da dimensão dos clientes ou a disponibilidade de informação (inputs do modelo), que importam relevar e que colocam em causa a aplicabilidade do modelo a outra realidade, que não aquela para a qual o modelo foi originalmente criado.

É sobretudo por motivos de normalização de procedimentos e de aproveitamento dos conhecimentos e da experiência obtidos em outra empresa do Grupo, que se dispensa este capítulo à verificação da aplicabilidade do modelo existente, à comercializadora que labora no mercado nacional.

Na realidade tratam-se de dois modelos distintos que utilizam os mesmos indicadores mas diferem nas respetivas ponderações.

O modelo identificado como “Peso Corporate”, aplica-se a empresas com volume de negócios anual superior a 100 milhões de euros, e o modelo para “Grandes Empresas” aplica-se a empresas com volume de negócios anual compreendido entre os 8 e os 100 milhões de euros.

Os modelos são constituídos por cinco componentes/módulos distintos:

a) Módulo Quantitativo ou Financeiro

O módulo quantitativo ou financeiro traduz-se na combinação de treze variáveis ou rácios financeiros, ponderados de acordo com o impacto que produzem na variável dependente, ou seja, na probabilidade da empresa entrar em incumprimento.

Quadro 3 – Modelo Atual - Rácios Económicos e Financeiros

Grupo	Rácio	Peso Corporate	Peso Grandes Empresas
Rendibilidade e Eficiência	Rendibilidade do Ativo - ROA	13%	12%
	Resultados Antes de Impostos (RAI)	3%	3%
	Rendibilidade do Capital Próprio	6%	6%
	Margem Bruta das Vendas	6%	6%
	Rendibilidade Líquida das Vendas	4%	3%
	Remuneração dos Capitais Alheios - Cobertura	13%	12%
		45%	42%
Estrutura Financeira	Solvabilidade	8%	13%
	Autonomia Financeira	5%	7%
	Peso Endividamento Longo Prazo	3%	6%
	Estrutura Financeira	4%	7%
		20%	33%
Liquidez e Solvência	Liquidez Geral	4%	4%
	Liquidez Imediata	6%	6%
		10%	10%
Rotação e Custos	Vendas	25%	15%
		25%	15%

Fonte: Documentação disponibilizada pela empresa em estudo

☐ Rendibilidade do Ativo (Rendibilidade Económica)

Rácio	POC	IAS/SNC
Rendibilidade do Ativo (ROA)	$\frac{\text{Resultados antes de juros e impostos}}{\text{Ativo Total Líquido}}$	$\frac{\text{Resultados antes de juros e impostos}}{\text{Ativo Total Líquido}}$

☐ **Resultado Antes de Impostos**

Rácio	POC	IAS/SNC
Resultados Antes de Impostos	Resultados Operacionais + Resultados Financeiros + Resultados Extraordinários	Resultados Operacionais (antes de gastos de financiamento e impostos) + Juros e Rend. Similares Obtidos - Juros e Gastos Similares Suportados

☐ **Rendibilidade dos Capitais Próprios (Rendibilidade Financeira)**

Rácio	POC	IAS/SNC
Rendibilidade dos Capitais Próprios	$\frac{\text{Resultados Líquidos}}{\text{Capitais Próprios}}$	$\frac{\text{Resultados Líquidos}}{\text{Capitais Próprios}}$

☐ **Margem Bruta das Vendas**

Rácio	POC	IAS/SNC
Margem Bruta	$\frac{\text{Resultado Antes Impostos}}{\text{Total Vendas}}$	$\frac{\text{Resultado Antes Impostos}}{\text{Total Vendas}}$

☐ **Rendibilidade Líquida das Vendas**

Rácio	POC	IAS/SNC
Rendibilidade Líquida das Vendas	$\frac{\text{Resultado Líquido do Exercício}}{\text{Total Vendas}}$	$\frac{\text{Resultado Líquido do Exercício}}{\text{Total Vendas}}$

☐ **Remuneração dos Capitais Alheios**

Rácio	POC	IAS/SNC
Remuneração Capitais Alheios	$\frac{\text{Encargos Financeiros}}{\text{Total Passivo}}$	$\frac{\text{Encargos Financeiros}}{\text{Total Passivo}}$

☐ **Garantia**

Rácio	POC	IAS/SNC
Garantia	$\frac{\text{Ativo Total}}{\text{Passivo Total}}$	$\frac{\text{Ativo Total}}{\text{Passivo Total}}$

☐ **Autonomia Financeira**

Rácio	POC	IAS/SNC
Autonomia Financeira	$\frac{\text{Capitais Próprios}}{\text{Ativo Total Líquido}}$	$\frac{\text{Capitais Próprios}}{\text{Ativo Total Líquido}}$

☐ **Peso Endividamento de Longo Prazo no Passivo Total**

Rácio	POC	IAS/SNC
Peso Endividamento LP no Passivo Total	$\frac{\text{Dívidas Longo Prazo}}{\text{Passivo Total}}$	$\frac{\text{Dívidas Longo Prazo}}{\text{Passivo Total}}$

☐ **Estrutura Financeira**

Rácio	POC	IAS/SNC
Estrutura Financeira	$\frac{\text{Passivo Total}}{\text{Capitais Próprios}}$	$\frac{\text{Passivo Total}}{\text{Capitais Próprios}}$

☐ **Liquidez Geral**

Rácio	POC	IAS/SNC
Liquidez Geral	$\frac{\text{Ativo Circulante}^{(1)}}{\text{Passivo de Curto Prazo}^{(2)}}$	$\frac{\text{Ativo Corrente}}{\text{Passivo Corrente}}$

(1) Existências + Dívidas de terceiros de Curto Prazo + Disponibilidades

(2) Dívidas a terceiros de Curto Prazo

☐ **Liquidez Imediata**

Rácio	POC	IAS/SNC
Liquidez Imediata	$\frac{\text{Disponibilidades}}{\text{Passivo de Curto Prazo}^{(2)}}$	$\frac{\text{Meios Financeiros Líquidos}}{\text{Passivo Corrente}}$

(2) Dívidas a terceiros de Curto Prazo

☐ **Vendas Líquidas**

Rácio	POC	IAS/SNC
Vendas Líquidas	Vendas Líquidas	Vendas Líquidas

b) Módulo Qualitativo

Este módulo é composto por um conjunto de perguntas pré-definidas, que se dividem entre perguntas do analista e perguntas do gestor comercial.

Cada uma das perguntas tem associadas três respostas que foram caracterizadas como risco alto, médio ou baixo. Sendo que cada pergunta só admite uma resposta.

c) Módulo Setorial

Consiste numa análise ao setor de atividade a que a empresa pertence, sintetizada em dois tipos de rating: Rating CP (curto prazo) e rating MP (médio prazo).

O rating pode tomar uma das seguintes codificações, do melhor para o pior: A⁺, A, A⁻, B⁺, B, B⁻, C⁺, C, C⁻.

d) Módulo de Alertas

O módulo de alertas, consiste num conjunto de informação considerada crítica, que só por si, penaliza ou classifica um cliente empresarial com risco elevado.

Quadro 4– Variáveis Críticas

Alarme	Descrição	Tipo
Informação - Hábitos de Pagamento Internos	O alarme será activado (assignando um valor =1) quando o cliente disponha de uma dívida vencida superior a 3.000 € no período de 0-30 dias ou superior a 1.500€ num período de 30-60 dias e sempre que não se deva a questões operacionais	Penalizadora
Informação - Hábitos de Pagamento Externos	O alarme será activado quando o saldo devedor do cliente seja superior a 3.000€ ou conte com mais de dois pagamentos em atraso, correspondentes a um saldo significativo (a definir).	Penalizadora
Dissolução	Mostra se uma empresa se encontra em situação potencial de dissolução segundo a L.S.A (Cap. Próprios < 50% Capital Subscrito)	Killer
Insolvência e Falência	O alarme será activado no caso em que uma empresa entre em processo de insolvência (ficheiro de INFORMA)	Killer
Informação Económica -Financeira Desfasada	Controla a antiguidade do balanço que está a ser utilizado para analisar o cliente. No caso de superar os 18/24 meses de antiguidade, o cliente já não será avaliado, obrigando à solicitação de um novo	Killer
Reclamações Jurídicas	Informação Jurídica contra o cliente (ficheiro de INFORMA)	Killer

Fonte: Documentação disponibilizada pela empresa em estudo

Convém ressaltar, que a informação do módulo de alertas se encontra efetivamente contemplada ao nível dos critérios de definição da amostra de clientes a utilizar no processo de verificação da aplicabilidade do modelo.

e) Módulo de Ajustamento ao Ciclo (conjuntura macroeconómica)

Este módulo tem por objetivo ajustar a classificação obtida nos módulos anteriores ao ciclo económico esperado nos próximos meses. Para tal, utilizam-se variáveis macroeconómicas como seja o Produto Interno Bruto (PIB).

De seguida, apresenta-se o quadro com a escala de pontuação produzida pelos modelos e a respetiva correspondência com o scoring e classificação do risco de crédito do cliente.

Quadro 5 – Escala de Pontuação dos Modelos

		Escala Níveis Rating		S&P	
		Níveis	Límite Pontuação	Nível	PD
Risco Baixo		1	94	AAA	0,00%
		2	88	AA+	0,00%
		3	82	AA	0,00%
		4	76	AA-	0,02%
		5	70	A+	0,05%
		6	64	A	0,07%
		7	58	A-	0,06%
		8	52	BBB+	0,16%
		9	47	BBB	0,25%
		10	42	BBB-	0,33%
Risco Médio	Baixo	11	36	BB+	0,57%
		12	30	BB	0,86%
		13	24	BB-	1,54%
Risco Médio	Alto	14	18	B+	2,70%
		15	12	B	7,10%
		16	6	B-	10,11%
Risco Alto		17	0	CCC-C	26,29%
		18	0	CCC-C	26,29%
		19	0	CCC-C	26,29%
Falência		20	0	CCC-C	26,29%

Fonte: Documentação disponibilizada pela empresa em estudo

Convém esclarecer que o scoring final produzido pelo modelo, resulta na soma do scoring alcançado no módulo quantitativo, com uma correção setorial (módulo setorial) e uma correção qualitativa (módulo qualitativo), conforme demonstrado pela figura 1.

Figura 1- Scoring Final – Correção Setorial e Qualitativa

	Rating	Qualitativo									
		1	2	3	4	5	6	7	8	9	10
Quantitativo	1	0	0	0	0	0	0	0	1	1	1
	2	-1	0	0	0	0	0	0	1	1	1
	3	-1	-1	0	0	0	0	0	1	1	1
	4	-1	-1	-1	0	0	0	0	1	1	1
	5	-1	-1	-1	0	0	0	0	1	1	1
	6	-1	-1	-1	0	0	0	0	1	1	1
	7	-1	-1	-1	0	0	0	0	1	1	1
	8	-1	-1	-1	0	0	0	0	1	1	1
	9	-1	-1	-1	0	0	0	0	1	1	1
	10	-1	-1	-1	0	0	0	0	1	1	1
	11	-1	-1	-1	0	0	0	0	1	1	1
	12	-1	-1	-1	0	0	0	0	1	1	1
	13	-1	-1	-1	0	0	0	0	1	1	1
	14	-1	-1	-1	0	0	0	0	1	1	1
	15	-1	-1	-1	0	0	0	0	1	1	1
	16	-1	-1	-1	0	0	0	0	1	1	1
	17	-1	-1	-1	0	0	0	0	1	1	1

	Rating	Setorial									
		1	2	3	4	5	6	7	8	9	10
Quantitativo	1	0	0	0	0	0	1	1	1	2	2
	2	-1	-1	0	0	0	1	1	1	2	2
	3	-1	-1	0	0	0	1	1	1	2	2
	4	-1	-1	0	0	0	1	1	1	2	2
	5	-1	-1	0	0	0	1	1	1	2	2
	6	-1	-1	0	0	0	1	1	1	2	2
	7	-1	-1	0	0	0	1	1	1	2	2
	8	-1	-1	0	0	0	1	1	1	2	2
	9	-1	-1	0	0	0	1	1	1	2	2
	10	-1	-1	0	0	0	1	1	1	2	2
	11	-1	-1	0	0	0	1	1	1	2	2
	12	-1	-1	0	0	0	1	1	1	2	2
	13	-1	-1	0	0	0	1	1	1	2	2
	14	-1	-1	0	0	0	1	1	1	2	2
	15	-1	-1	0	0	0	1	1	1	2	2
	16	-1	-1	0	0	0	1	1	1	1	1
	17	-1	-1	0	0	0	0	0	0	0	0

Fonte: Documentação disponibilizada pela empresa em estudo

O módulo quantitativo é o que possui maior peso na determinação da qualidade do crédito dos clientes. É por isso que lhe é dado especial enfoque, e será exclusivamente em torno das variáveis que o constituem que vai ser desenvolvido todo o trabalho, conforme referido anteriormente.

Neste estudo, as variáveis qualitativas do modelo serão negligenciadas, certa de que a sua contribuição é meramente um indicador de reforço de um estado económico e financeiro que só por si deixa adivinhar a capacidade do cliente solver as suas dívidas.

Capítulo 3 - Verificação da Aplicabilidade do Modelo

O estudo para averiguar se os modelos referidos no capítulo anterior, na sua vertente quantitativa, são eficazes na classificação do crédito dos clientes empresariais da empresa que labora no mercado nacional, consiste na aplicação desses modelos a uma amostra de clientes.

Espera-se que os modelos consigam prever, com elevado grau de precisão, o incumprimento do cliente empresarial com um horizonte temporal de pelo menos um ano.

De acordo com o exposto e atendendo ao facto de as empresas serem obrigadas a disponibilizar a respetiva informação financeira anual até ao mês de Julho do ano seguinte, (o período de entrega este ano foi excecionalmente alargado), será utilizada a informação financeira dos clientes (Balanço e Demonstração de Resultados), referente ao ano de 2009.

A verificação da aplicabilidade do modelo, propriamente dita, será efetuada com base numa análise comparativa entre os resultados obtidos com a aplicação dos modelos à amostra e o estado atual desses clientes, ou seja, a verificação real de os mesmos se encontrarem ou não em estado de incumprimento, à data do estudo (Julho 2011).

3.1. Critérios Para Classificação do Risco dos Clientes

Os clientes foram sujeitos a uma classificação de risco de crédito, baseada nos critérios que a seguir se apresentam, bem como, na simples constatação do seu estado atual de incumprimento. Os clientes podem assumir dois tipos de classificação: “Sem Risco” e “Com Risco”.

■ Clientes “Sem Risco”

O cliente não possui qualquer tipo de registo na base de dados de contencioso ou de dívida vencida.

Numa base de alargamento da dimensão da amostra, foram também considerados alguns clientes com dívida vencida inferior a 6 meses, mas que possuem registo de reclamações em aberto enquadradas no âmbito da faturação (leitura de consumos erradas, dados do cliente incorretos...). Nestes casos, o incumprimento não pode ser imputado ao cliente.

■ Clientes “Com Risco”

O cliente possui registo na base de dados de contencioso e/ou dívida vencida, não tendo qualquer registo de reclamação ou pedido associado.

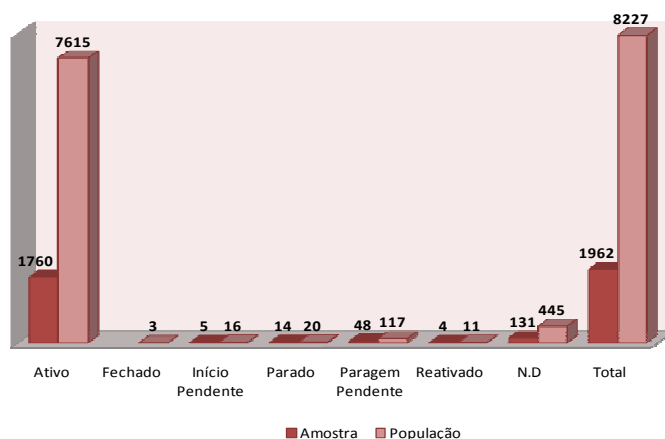
A definição dos critérios tem por base a noção mais básica de incumprimento, que se traduz no registo do não pagamento, por parte do cliente, uma vez ultrapassado o prazo de vencimento para realizar o mesmo. A esta noção, e de modo a obter uma classificação mais fidedigna foram considerados alguns aspectos, relacionados com a vontade alheia do cliente para que o incumprimento ocorresse.

3.2. Definição e Caracterização da Amostra

A base de dados de clientes teve origem num relatório extraído do sistema comercial da área de negócio B2B em Maio de 2011, tendo sido considerados todos os clientes registados na mesma, independentemente do seu status.

Importa esclarecer, que embora um dos termos de comparação para avaliar a aplicabilidade dos modelos seja o estado atual dos clientes, interessariam à primeira vista, somente os clientes ativos na carteira, à data do estudo, no entanto, numa ótica de alargamento da dimensão da amostra e também de considerar os clientes que possam ter entrado em incumprimento no período compreendido entre 2009 e 2011, e por isso possam já não estar ativos à data do estudo, foi considerada toda a base de dados de clientes.

Gráfico 1 – Status de Clientes



Dos 8227 clientes (População) com cadastro na base de dados, foram selecionados 1962 (Amostra), encontrando-se a maior parte deles ativos (1760). Foram também considerados para a amostra, alguns clientes com status “parado”, “pendente” ou “reativado”, por terem entrado em incumprimento entre os anos de 2009 e 2010. Foram ainda selecionados 131 clientes sem status definido.

De forma a obter uma amostra representativa da carteira de clientes B2B da empresa, foram definidos os seguintes critérios:

☐ Respeitar a representatividade de todos os segmentos de negócio que compõem a População, tendo-se enveredado por um processo de seleção da amostra estratificado;

Os clientes da empresa encontram-se segmentados de acordo com o seguinte:

Quadro 6 – Definição dos Segmentos

Segmentos	Consumo Anual Estimado de Energia
Diamante	>= 20 GWh
Platina	2 - 20 GWh
Ouro	0,75 - 2 GWh
Prata	0,2 - 0,75 GWh

Fonte: Empresa em Estudo

Gráfico 2 – População Clientes/Segmento

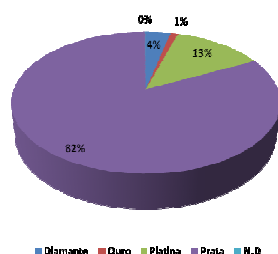
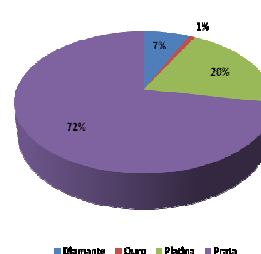


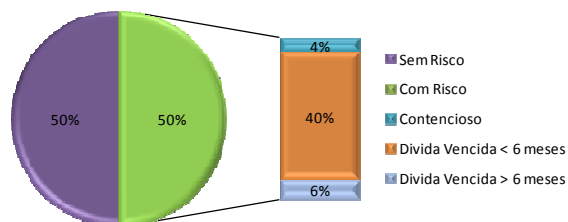
Gráfico 3– Amostra Clientes/Segmento



□ Igual representatividade de clientes “Com Risco” e “Sem Risco”, ou seja, seleção do mesmo número de clientes “Com Risco” e “Sem Risco”, por cada segmento.

Este método de emparelhamento da amostra baseia-se, nomeadamente, nos trabalhos de Beaver (1966), Altman(1968) e Sanvicente e Minardi(1998).

Gráfico 4 – Amostra – Classificação de Clientes



□ Respeitar a representatividade da localização geográfica e atividade económica da População;

Gráfico 5 – Percentagem de Clientes por Localização Geográfica

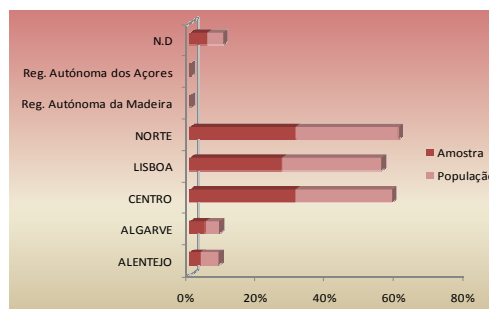
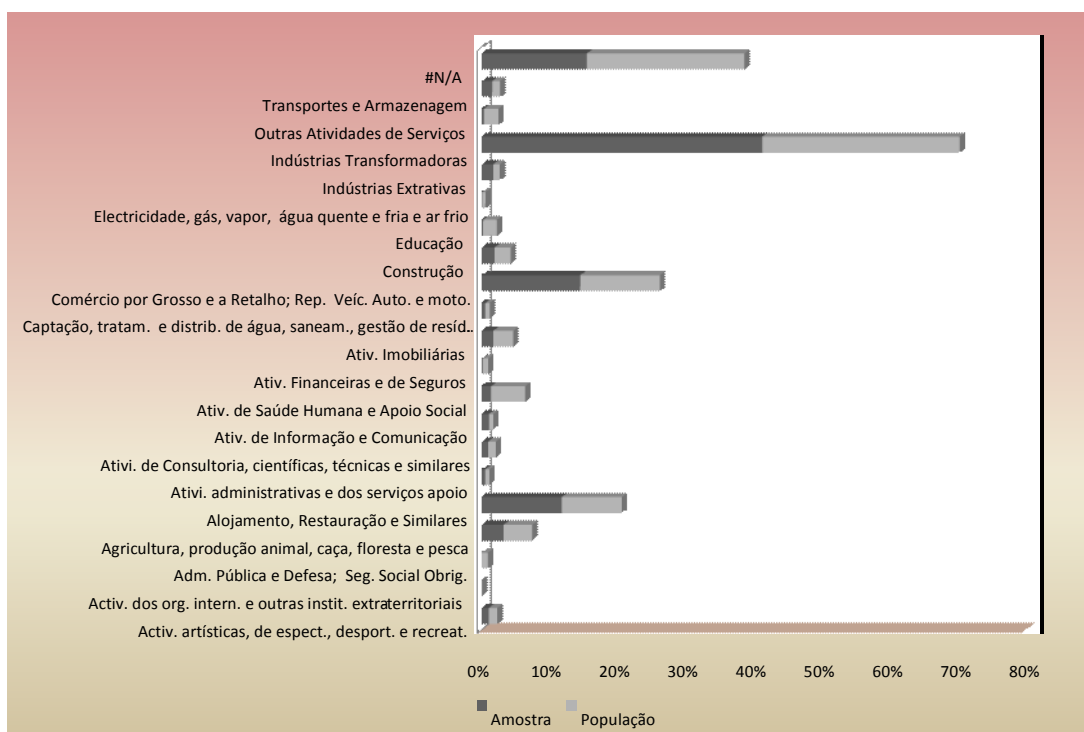


Gráfico 6 – Percentagem de Clientes por Atividade Económica



☐ Clientes com data de ativação igual ou posterior a 01.01.2009.

☐ Clientes com histórico de faturação na empresa superior a 6 meses, de modo a poder aferir sobre o seu comportamento de pagamentos.

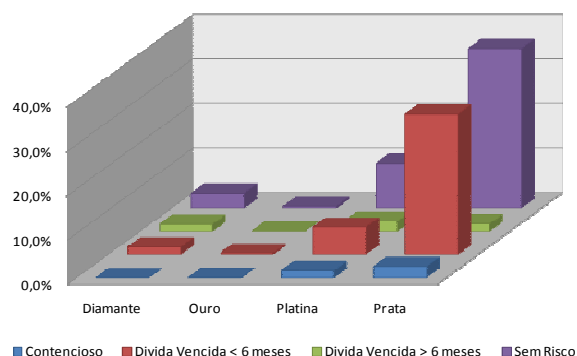
☐ Clientes do setor estado: Considerar somente o setor empresarial do estado.

O setor Estado representa atualmente uma fatia interessante da carteira ativa de clientes B2B, tratando-se essencialmente de organismos públicos ligados à administração central, regional e local e algumas empresas públicas. Pela sua natureza e porque a respetiva informação financeira tem uma leitura distinta das demais empresas do setor empresarial, optou-se pela não consideração daqueles organismos.

Pelo mesmo motivo, foram também excluídas as IPSS's (Instituições Particulares de Solidariedade Social) e outras entidades sem fins lucrativos, as quais funcionam na sua maioria com apoios estatais.

A composição da amostra, nesta fase, é a que a seguir se apresenta e evidencia a representatividade das características fundamentais da população.

Gráfico 7 – Amostra – Classificação de Clientes Por Segmento



Conforme referido no capítulo anterior, existem dois modelos que vão ser testados e que podem ser apelidados de “Modelo 1” e “Modelo 2”, sendo que o primeiro será aplicado às empresas que possuem uma faturação anual compreendida entre os 8 e os 100 milhões de euros (“Grandes Empresas”), e o segundo destina-se às empresas com faturação superior a 100 milhões (Peso Corporate).

Após a seleção da amostra de 1962 clientes, foi solicitado a uma entidade externa, o fornecimento dos respetivos dados financeiros referentes ao ano de 2009.

A disponibilização desta informação, permitiu verificar quais os clientes que possuíram uma faturação anual compreendida entre os 8 e os 100 milhões de euros (327 clientes), e quais os que obtiveram valores superiores a 100 milhões de euros (48 clientes). A amostra final será assim, constituída por 375 clientes.

Quadro 7– N.º de Clientes Por Intervalo de Volume de Negócios

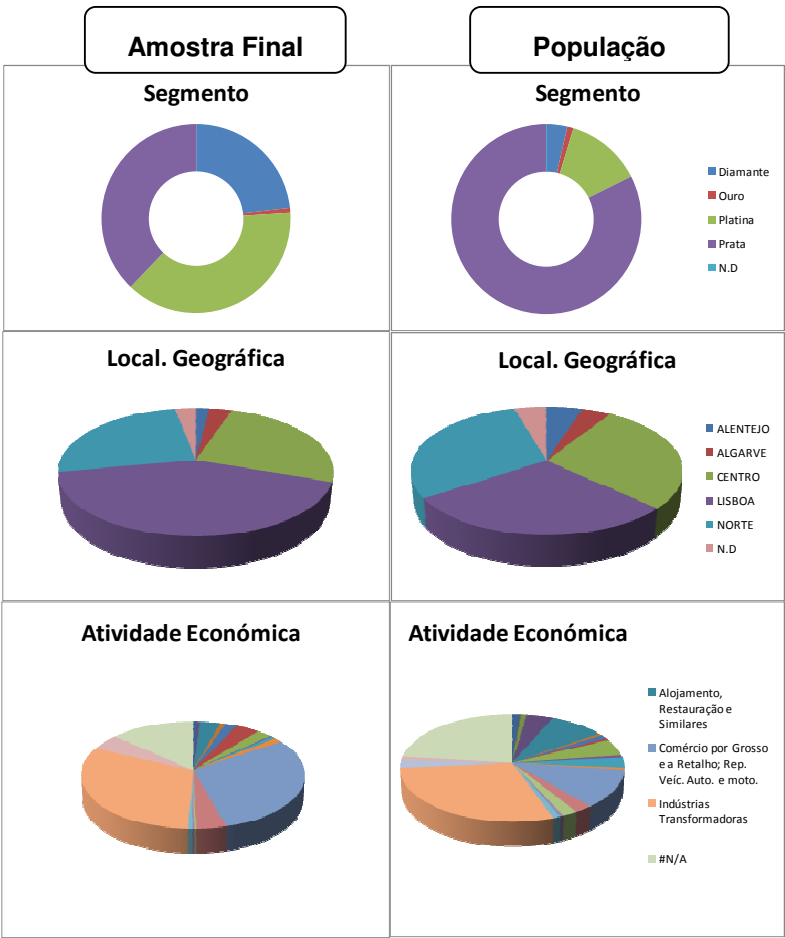
Faturação Anual	Representatividade de Clientes na Amostra	
	Número	Percentagem
< 8 Milhões	1587	81%
8 - 100 Milhões	327	17%
> 100 Milhões	48	2%
Total	1962	100%

Convém ressaltar, que embora se soubesse à partida que os modelos só seriam testados em empresas com faturação anual dentro dos intervalos referidos (clientes com volume de negócios superiores a 8 milhões de euros), não foi possível numa fase anterior restringir a seleção da amostra a clientes com aquela característica, uma vez que: A empresa não

dispõe de informação sobre o volume de negócios de cada cliente, sendo que a informação financeira dos clientes é adquirida a entidades externas, representando um custo significativo se adquirida para toda a carteira.

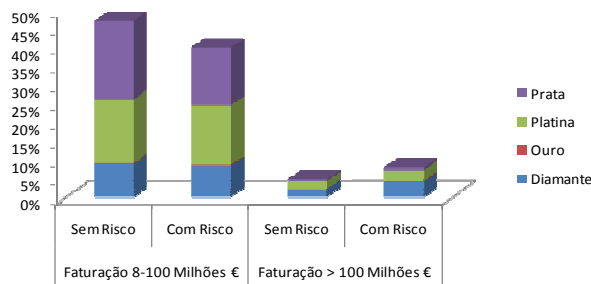
Veja-se de seguida a figura 2, que permite comparar as características base da população e da amostra final de clientes, verificando-se que os critérios de seleção da amostra continuam a ser cumpridos.

Figura 2 – Caracterização População vs Amostra Final



Veja-se adicionalmente a classificação do risco da amostra por segmento e intervalo de faturação:

Gráfico 8 – Composição da Amostra Final de Clientes



Em linha de conta com o que tem vindo a ser descrito neste capítulo, verifica-se que a amostra final atende à representatividade real de cada segmento de clientes e permite testar equitativamente o número de casos de “Clientes Com Risco” e “Clientes Sem Risco”. Embora não se verifique uma repartição equilibrada dos clientes pelos intervalos de faturação, acredita-se que a representatividade na carteira real de clientes da empresa, dos clientes com níveis de faturação superiores a 100 milhões de euros é bastante inferior à dos clientes com faturação compreendida entre os 8 e os 100 milhões de euros.

A inferência através da informação do segmento do cliente, sobre o volume de negócios do mesmo pode ser uma falácia, pois como bem se vê, existem clientes do setor prata (consumos de energia inferiores) com faturação acima dos 100 milhões de euros.

3.3. Aplicação dos Modelos à Amostra

3.3.1. Resultados Obtidos

Para cada empresa da amostra (375) foram calculados os rácios económicos e financeiros que definem os modelos, estes cálculos foram realizados à luz do POC (Plano Oficial de Contabilidade), uma vez que se pretende obter conclusões, atendendo aos critérios originalmente utilizados na definição dos modelos que vão ser testados. Porém, uma breve análise comparativa entre as fórmulas de cálculo dos rácios à luz do POC e do SNC (Sistema de Normalização Contabilística), com base no capítulo 2, permite concluir que não se sentem diferenças significativas no valor dos rácios.

Num passo seguinte, através da aplicação das respetivas ponderações aos rácios, e atendendo ao volume de negócios de cada empresa, obteve-se a pontuação quantitativa

(PQ) do modelo para cada uma das empresas, que de acordo com a tabela de escala de pontuação dos modelos (Quadro 5 da p. 39) foi convertido para um scoring e classificação de risco de crédito de cada empresa.

Relembrem-se agora os modelos:

• **Modelo 1 – Grandes Empresas**

$$PQ = 0,12 * X_1 + 0,03 * X_2 + 0,06 * X_3 + 0,06 * X_4 + 0,03 * X_5 + 0,12 * X_6 + 0,13 * X_7 + 0,07 * X_8 + 0,06 * X_9 + 0,07 * X_{10} + 0,04 * X_{11} + 0,06 * X_{12} + 0,15 * X_{13}$$

• **Modelo 2 – Peso Corporate**

$$PQ = 0,13 * X_1 + 0,03 * X_2 + 0,06 * X_3 + 0,06 * X_4 + 0,04 * X_5 + 0,13 * X_6 + 0,08 * X_7 + 0,05 * X_8 + 0,03 * X_9 + 0,04 * X_{10} + 0,04 * X_{11} + 0,06 * X_{12} + 0,25 * X_{13}$$

Sendo,

PQ = Pontuação Quantitativa

X_8 = Autonomia Financeira

X_1 = Rendibilidade do Ativo

X_9 = Peso Endiv. Longo Prazo

X_2 = Ln Resultado Líquido do Exercício

X_{10} = Estrutura Financeira

X_3 = Rendibilidade dos Capitais Próprios

X_{11} = Liquidez Geral

X_4 = Margem Bruta das Vendas

X_{12} = Liquidez Imediata

X_5 = Rendibilidade Líquida das Vendas

X_{13} = Ln Vendas

X_6 = Remuneração dos Capitais Alheios

X_7 = Garantia

A aplicação, quer de um modelo, quer de outro, resultaram na classificação de todas as empresas da amostra como tendo “Risco Alto” ou “Risco Médio Alto”.

Quadro 8 – Modelo 1 – Matriz de Classificação

Modelo 1 - Grandes Empresas		
Observado	Previsto	
	Com Risco	Sem Risco
Com Risco	100%	0%
Sem Risco	100%	0%

O modelo apresenta uma taxa de acerto de 100%, no caso dos clientes “Com Risco”, ou seja, aqueles que entraram em incumprimento, no entanto, não acertou em nenhum dos casos classificados como “Sem Risco”, apesar da amostra incluir 54% casos com esta classificação.

Quadro 9 – Modelo 2 – Matriz Classificação

Modelo 2 - Peso Corporate		
Observado	Previsto	
	Com Risco	Sem Risco
Com Risco	100%	0%
Sem Risco	100%	0%

À semelhança do modelo 1, este modelo também classifica todas as empresas da amostra como “Com Risco”, apesar de 38% da amostra ser constituída por casos de clientes “Sem Risco”.

3.3.2. Conclusão

Como facilmente se compreende, a percentagem de casos com previsões correctas não é um critério suficiente para avaliar as capacidades previsionais de um modelo. De facto, torna-se necessário tomar em consideração, na avaliação do modelo, os custos e os benefícios de classificar bem ou mal uma empresa. Os erros de classificação podem ser de dois tipos, e designam-se respetivamente por erros de tipo I e erros de tipo II.

Quadro 10 - Tipos de Erro

Observado	Previsto	
	Com Risco	Sem Risco
Com Risco	Correto	Erro Tipo I
Sem Risco	Erro Tipo II	Correto

Os custos associados a cada um destes tipos de erro não é igual. Altman et al. (1977), num modelo com um âmbito diferente do presente estudo, estimaram que o custo de um erro de tipo I seria 35 vezes superior ao custo de um erro de tipo II. Isto é, seria bastante mais grave estimar que uma empresa não tem risco quando afinal vai entrar em incumprimento, do que prever que a empresa vai entrar em incumprimento, quando afinal não tem risco associado. Esta situação pode ser exemplificada do seguinte modo:

- a) No caso de ocorrer um erro de tipo I significava que o credor iria contratar um cliente que entrará em incumprimento podendo vir a perder os valores de energia faturados, os juros, e até mesmo, incorrer em custos extraordinários de reforço dos serviços de cobrança e jurídicos.
- b) O erro tipo II significa não contratar um potencial cliente que, afinal, tinha condições para cumprir com as suas obrigações, perdendo assim a margem que o cliente potenciava.

Os resultados obtidos revelam que os modelos, tal como foram criados para ser aplicados numa outra geografia e a um universo de clientes com outras características e dimensões, não se aplicam à amostra de clientes e por inferência, não se aplicam ao universo de clientes em causa. Sendo justamente ao nível do erro tipo I que ambos os modelos mais pecam.

Para além das razões já evidenciadas, outros motivos, de ordem mais técnica, poderão ser evocados para justificar a não aplicabilidade do modelo, nomeadamente o facto de não se poder fazer uma leitura exatamente igual dos rácios utilizados nas duas geografias, uma vez que as regras de contabilização utilizadas são diferentes nos dois países.

Outro aspecto, prende-se com a segmentação da carteira utilizada para a aplicação de um e de outro modelo. O facto de em Portugal o tecido empresarial se revestir de uma outra dimensão e estrutura setorial comparativamente ao do país onde os modelos são originalmente aplicados, leva a colocar em discussão a definição de novos modelos, através de um ajustamento da segmentação da carteira e muito provavelmente dos rácios económicos e financeiros selecionados.

Capítulo 4 - Definição de Modelos Alternativos

Neste capítulo, procura-se definir e estimar um ou vários modelos alternativos aos analisados no capítulo anterior. Uma vez que estes se revelaram inadequados, será necessário prosseguir com o levantamento de requisitos e a definição de novos critérios, que respondam à realidade da carteira real de clientes.

Relembre-se que o objetivo do trabalho cinge-se à procura de um modelo de avaliação do risco, baseado na análise financeira do cliente por comparação com os modelos testados.

Um dos primeiros passos para a estimação do modelo, passa pela definição das variáveis que melhor consigam explicar o fenómeno que se pretende prever.

Neste caso, procura-se encontrar as variáveis quantitativas (rácios económicos e financeiros) que melhor consigam denunciar o incumprimento de clientes, com um período de antecipação de pelo menos um ano.

Outro aspecto, não menos importante, prende-se com a definição da segmentação da carteira.

4.1. Segmentação da Amostra

A amostra de clientes definida, representativa da população, conforme demonstrado anteriormente, é constituída por clientes com diferentes dimensões, isto é, empresas que registam anualmente, diferentes níveis de volume de negócios.

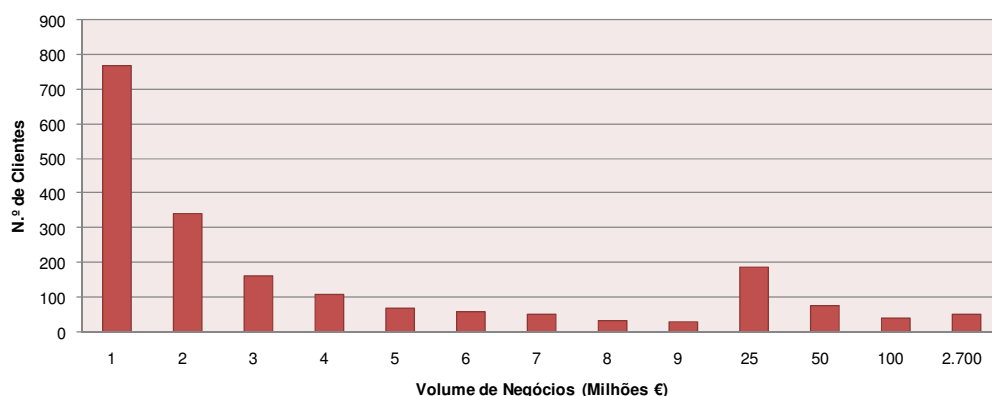
Embora o significado dos rácios seja universal, também é verdade que os mesmos podem assumir diferente interpretação e importância, consoante a atividade e a dimensão da empresa em causa, tome-se como exemplo, o caso flagrante da estrutura de endividamento de um banco e de uma empresa fabril. Este é um dos motivos, pelos quais se decidiu manter o mesmo critério de segmentação de clientes, utilizado na outra empresa do Grupo, e que se baseia no volume de negócios (VN) anual do cliente.

Considera-se assim, que este critério permitirá obter uma modelização mais ajustada e resultados mais precisos.

Logo à partida, descartou-se a hipótese de utilização de segmentação da carteira utilizada na outra empresa do Grupo (Grandes Empresas – VN entre 8 e 100 milhões de euros e Empresas com Peso Corporate – VN superior a 100 milhões), uma vez que se trata de um tecido empresarial com características geográficas, setoriais e dimensão diferentes, face às empresas nacionais. A própria constituição das duas carteiras de clientes é diferente em termos de dimensão dos clientes.

O processo de delineamento das “zonas” de segmentação, ou seja, do ou dos valores de VN que devem estabelecer a aplicabilidade dos modelos, começou por ter em atenção a distribuição de frequência de clientes (amostra inicial de 1962 clientes), por nível de VN.

Gráfico 9 – Frequência de Clientes Por Intervalo de VN



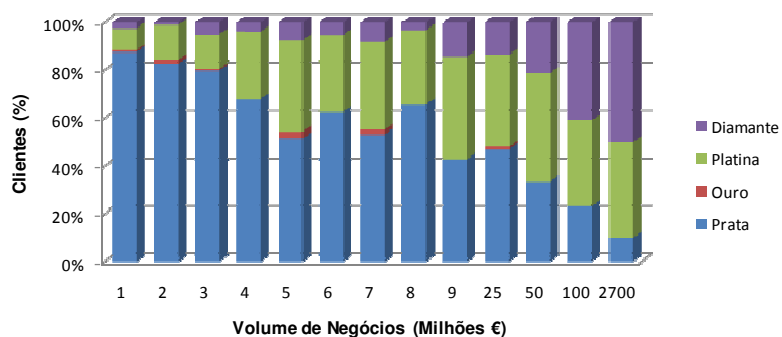
Através do gráfico, é possível constatar que existe uma grande concentração de clientes no intervalo de VN compreendido entre os zero e os dois milhões de euros (56%), ou seja, mais de metade da amostra representativa da população regista VN inferiores a dois milhões de euros, sendo que 39% dos clientes não ultrapassa o um milhão de euros.

Outro aspecto que pesou na definição da segmentação, tem a ver com as características do setor empresarial nacional. Segundo dados do INE relativos a 2008, a dimensão média das empresas portuguesas é muito reduzida, empregam em média, 8,6 trabalhadores e registam em média, cerca de 993,3 mil euros de vendas. No caso das pequenas empresas, os valores descem para 6,2 trabalhadores e para 576 mil euros respetivamente.

Embora a completa liberalização do setor energético aumente o número de clientes exigíveis, permitindo o aumento da representatividade de grandes empresas na carteira, acredita-se que existirá sempre uma grande fatia de pequenas empresas clientes, que interessa conseguir classificar eficazmente, em termos de risco de crédito.

Foi ainda elaborada, uma análise sobre a representatividade de cada segmento de negócio por intervalo de volume de negócios dos clientes, de modo a poder estabelecer algum tipo de ligação entre os modelos de risco de crédito a estimar e a segmentação do negócio.

Gráfico 10 – Representatividade dos Segmentos Por Intervalo de VN



A análise simultânea da representatividade dos segmentos de negócio, e da mesma em cada intervalo de VN, permite concluir que apesar de se verificar uma predominância de clientes do segmento prata (consumos mais pequenos) no intervalo das empresas com volume de negócios mais pequenos até 7/8 milhões de euros, também se verifica uma representatividade significativa deste segmento em intervalos de faturação superiores. Esta análise aplica-se igualmente aos restantes segmentos. Por exemplo, os clientes dos segmentos Platina e Diamante (grandes consumos), distribuem-se quase equitativamente pelos vários intervalos definidos.

O comportamento desta distribuição de frequências do número de clientes por segmento e nível de volume de negócios, permite refutar completamente, a ideia de que clientes com grandes consumos são grandes empresas e vice versa.

Assim, concluí-se que não faz muito sentido estabelecer uma ligação entre o que deverá ser a segmentação da carteira para a análise do risco de crédito, e a segmentação do negócio.

Como é óbvio, para a empresa comercializadora de energia, um cliente que possua um maior peso na carteira, porque pertence a um segmento de negócio mais elevado, representará naturalmente, uma maior perda se entrar em incumprimento, do que se se tratar de um cliente que integre um segmento mais baixo e tiver a mesma sorte, no entanto, esta é uma outra questão. Para a empresa, interessa numa fase inicial, perceber somente, qual é a capacidade do cliente cumprir as suas obrigações, posteriormente, diante da classificação do risco do cliente obtida, o gestor decidirá sobre a forma de mitigar esse risco, e este, é um aspecto que cai fora do âmbito do presente trabalho.

De acordo com o exposto, decidiu-se, numa primeira fase, estabelecer duas possíveis “zonas” de segmentação: A primeira estabelece dois modelos diferentes para empresas com volume de negócios inferiores e superiores a um milhão de euros, e a segunda, dois modelos para aplicação em empresas com volume de negócios superior ou inferior a dois milhões de euros.

4.2. Seleção das Variáveis

Uma vez que não existe nenhuma teoria consolidada sobre quais os melhores indicadores económicos e financeiros capazes de prever o default de uma empresa, optou-se pela seleção dos rácios financeiros utilizados nos trabalhos desenvolvidos por outros autores, nomeadamente, Blum (1974), Beaver (1966), Altman (1968, 1974, 1984), Edminster (1972), Deakin (1972), Taffler (1982), Stein e Ziegler (1984), Altman & Sabato (2005) tendo-se adicionalmente, considerado os rácios dos modelos anteriormente testados, de modo a alargar o leque de possíveis combinações.

Os rácios económicos e financeiros identificados compreendem indicadores de liquidez, de rendibilidade, de funcionamento e de alavancagem.

Os rácios de rendibilidade são indicadores que avaliam os resultados gerados pela empresa. Os indicadores de liquidez permitem medir a capacidade de pagamento, comparando nomeadamente o ativo realizável momentaneamente com as exigibilidades. Por seu lado, os indicadores de alavancagem, medem a proporção de recursos próprios em relação aos recursos de terceiros na estrutura de capital, e a capacidade de cobertura da empresa para pagar os seus encargos financeiros. Os rácios de funcionamento

permitem aferir sobre a aplicação dos recursos, nomeadamente, a cobertura dos ativos e passivo pelas vendas.

As variáveis seleccionadas são as que a seguir se apresentam:

Quadro 11 – Variáveis Seleccionadas

Tipologia do Rácio	Variável	Rácios	Fórmula de Cálculo
Rendibilidade	A1	ROA - Rendibilidade do Ativo	Resultado Antes de Impostos/Ativo Total
	A2	ROE - Rendibilidade Capital Próprio	Res. Líquidos/Capital Próprio
	A3	Rendibilidade Acumulada	Resultados Retidos/Ativo Total
	A4	Margem Bruta das Vendas	Resultado Antes de Impostos/Vendas
	A5	Rend. Líquida das Vendas	Res. Líquidos/Vendas
Liquidez	A6	Liquidez Imediata	Meios Financeiros Líquidos/Dívidas a Terceiros C/P
	A7	Rácio Fundo de Maneio	Fundo de Maneio/Ativo Total
	A8	Meios Financeiros Líquidos sobre Ativo	Meios Financeiros Líquidos/Dívidas a Terceiros C/P
	A9	Liquidez Geral	Ativo Circulante/Passivo Circulante
	A10	Vendas Por Meios Financeiros Líquidos	Vendas/MFL
	A11	Fundo de Maneio Sobre Capital Próprio	Fundo de Maneio/Capital Próprio
Alavancagem	A12	Cobertura dos Custos Financeiros	Resultados Operacionais/Gastos Financeiros
	A13	Dívidas a Terceiros C/P sobre Capitais Próprios	Dívidas a Terceiros C/P/Total Capital Próprio
	A14	Remuneração Capitais Alheios	Gastos Financeiros/Passivo Total
	A15	Peso End. LP	Div. A Terc. LP/Passivo Total
	A16	Estrutura Financeira/Debt to equity ratio	Total Passivo/Capital Próprio
	A17	Endividamento	Passivo Total/Ativo Total
	A18	Estrutura do Cap Permanente	Cap. Alheio m/l prazo/Cap. Próprio
	A19	Cobertura do Ativo Fixo Tangível	Ativo Fixo Tangível/Cap. Permanentes
	A20	Cobertura do Ativo Por Resultados Transitados	Resultados Transitados/ Ativo Total
	A21	Estrutura do Endividamento	Passivo Circulante/Passivo Total
Solvabilidade	A22	Autonomia Financeira	Total Capital Próprio/Passivo Total
	A23	Solvabilidade	Situação Líquida/Passivo Total
Funcionamento	A24	Rotação de Fundo de Maneio	Vendas/Fundo de Maneio
	A25	Rotação do Capital Próprio	Vendas/Capital Próprio
	A26	Rotação do Ativo	Vendas/Ativo Total
	A27	Vendas Por Contas a Receber	Vendas/Div. De Terceiros
	A28	Rotação das Existências	Vendas/Existências
Outros Indicadores: B1 - Ativos Tangíveis B2 - Vendas B3 - Res. Antes Impostos			

O motivo pelo qual se seleccionaram vários rácios dentro da mesma tipologia, prende-se com o facto de se estar a analisar empresas pertencentes a diferentes setores de atividade e dimensão, sendo necessário contemplar o maior número possível de indicadores sobre a performance financeira e económica de uma empresa. Na realidade, dentro de cada tipologia de rácios, consegue-se distinguir uma certa heterogeneidade de indicadores que poderão ser mais ou menos explicativos do incumprimento, consoante a estrutura económica e financeira da empresa em causa.

De referir, as duas variáveis (A8 e A13), sugeridas num trabalho de Altman & Sabato (2005), relativo à definição de um modelo de avaliação do risco de crédito para PME's americanas, com vendas anuais inferiores a 65 milhões de dólares.

Uma vez que a carteira real de clientes possui uma percentagem importante de pequenas e médias empresas, parece razoável que se dê especial atenção a esta fatia da população.

4.3. Metodologia

Desde o trabalho pioneiro de Beaver (1966) diferentes métodos têm sido utilizados na previsão do incumprimento: Análise Discriminante, logit, probit, redes neuronais, indução de regras e árvores de decisão, algoritmos genéticos, conjuntos aproximados, entre outros modelos. O objetivo do presente trabalho não foi efetuar uma comparação profunda entre os métodos mais utilizados, mas sim utilizar a metodologia mais prática e com melhores resultados. O trabalho de revisão de literatura, permite concluir que a capacidade de previsão dos modelos é em geral similar, sendo que a maioria dos investigadores tem utilizado a análise discriminante ou o logit, apresentando este modelo algumas vantagens sobre o primeiro. Deste modo, a decisão sobre o modelo a utilizar na estimação de modelos alternativos de previsão do risco de crédito de clientes, recaiu sobre o modelo logit.

4.3.1. Modelo Logit

O modelo "logit" baseia-se na função logística que surgiu em 1845 ligada a problemas de crescimento demográfico, (Meadows et al.,1972).

Trata-se de um modelo econométrico de seleção qualitativa, uma vez que gera respostas de procedimentos qualitativos, como é o caso se uma empresa entra ou não em default. A este propósito, cite-se a obra de Cramer (1991) onde é efetuada uma introdução aos modelos "logit" tratando-os como um "standard" para modelos de escolha qualitativa.

Segundo Hair (1999), uma representação dos modelos, em que a variável dependente é dicotómica é a que pressupõe que a variável Y , (com $Y=1$ ou $Y=0$), é apenas uma indicação observável de uma variável não observável Y^* (variável latente) tal que:

$$Y_i^* = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i = X_i \beta + u_i$$

em que é especificada uma regra de determinação de Y em função de Y*.

Essa regra é, tipicamente, da seguinte forma:

$$Y_i = \begin{cases} 1, & \text{Se } Y^* \geq 0 \\ 0, & \text{Se } Y^* < 0 \end{cases}$$

onde a escolha da constante 0 como limiar da separação entre Y=0 e Y=1 é, essencialmente arbitrária.

Assim, pode escrever-se:

$$\begin{aligned} \text{Prob}(Y_i = 1) &= \text{Prob}(Y_i^* > 0) = \\ &= \text{Prob}(\beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i) > 0 = \\ &= \text{Prob}(u_i > -X_i \beta) \end{aligned}$$

e por conseguinte,

$$\text{Prob}(Y_i = 0) = \text{Prob}(u_i \leq -X_i \beta).$$

Então, sendo u_i uma variável aleatória com função de distribuição $F(\cdot)$, vem que:

$$\text{Prob}(Y_i = 0) = F(-X_i \beta)$$

$$\text{Prob}(Y_i = 1) = 1 - F(-X_i \beta).$$

No caso do modelo logit, em que a forma funcional de $F(\cdot)$ é a que corresponde à distribuição logística, vem:

$$E(Y_i) = 0 [1 - F(X_i \beta)] + 1 F(X_i \beta),$$

ou seja,

$$E(Y_i) = F(X_i \beta).$$

Portanto,

$$E(Y) = \text{Prob}(Y = 1) = \frac{1}{1 + e^{-x}}$$

onde,

$$\text{Prob}(Y_i = 1) = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_{21} + \dots + \beta_k X_{ki})}}$$

Para o mesmo autor, a regressão logística binária corresponde a uma regressão aplicada a uma variável dependente dicotómica, onde esta variável dependente não corresponde aos valores dos dados, mas representa a probabilidade do evento em análise vir a acontecer.

Supõe-se que o logaritmo da razão entre as probabilidades de ocorrência e não ocorrência do evento é linear. Como o objetivo do estudo é a probabilidade de acontecer determinado evento, a equação logística pode ser transformada numa equação de probabilidade assumindo assim a fórmula atrás apresentada.

Portanto, y é uma variável binária que, por convenção, assume os valores 0 ou 1 que correspondem à definição das sub-populações, sendo \mathbf{x} um vector coluna de dimensão $(p+1)$ (por forma a incluir as p variáveis que se consideraram na secção anterior mais um termo independente) de variáveis exógenas, β é um vector de parâmetros desconhecidos e não observáveis que se pretende estimar.

Na regressão logística, uma vez que os erros do modelo seguem uma distribuição de Bernoulli, o método para estimar os coeficientes β_0, \dots, β_n é o método da máxima verosimilhança, o qual procura determinar os valores dos parâmetros β , que maximizam a probabilidade de se obter o conjunto dos dados observados (Hosmer e Lemeshow, 1998).

Na análise logística existem os métodos *enter*, *backward* e *forward*, os quais diferem entre si pela forma como seleccionam as variáveis que irão formar o modelo final.

A regressão logística minimiza o número de variáveis dependentes, de forma que o modelo final, possua mais aplicabilidade, uma vez que quantas mais variáveis existirem, mais dependente de dados fica o modelo.

O método *enter* caracteriza-se por incorporar ao modelo todas as variáveis em estudo (deverá assim ser utilizado quando se tem certeza de que todas as variáveis são necessárias para estimar os parâmetros do modelo).

O método *backward*, caracteriza-se por introduzir no modelo todas as variáveis e com o decorrer do processo vai sendo eliminada uma variável de cada vez, de forma que quando o processo de eliminação termina, temos definido o modelo final. Este processo ocorre assim por etapas, onde em cada etapa a variável de pior desempenho é eliminada, comparando-se o modelo completo com o modelo reduzido, resultante da eliminação de variáveis (Charnet *et al.*, 1999). O modelo resultante final corresponde assim àquele, com as variáveis que melhor caracterizam o objeto do estudo.

O método *forward*, caracteriza-se por considerar a variável de maior coeficiente de correlação na amostra, observado com a variável resposta. Em cada etapa uma nova variável pode ser adicionada de forma a construir o modelo final. Este processo ocorre assim por etapas, onde em cada etapa uma nova variável é acrescentada, comparando-se o modelo reduzido com o modelo em que uma nova variável é acrescentada. O processo decorre enquanto se verificar que com a inclusão de uma variável se obtém um modelo com melhor desempenho que o anterior. Quando o processo de inclusão de variáveis termina, é porque não faz sentido a inclusão de mais nenhuma variável e as variáveis incluídas até essa data correspondem às que melhor caracterizam o objeto do estudo (Charnet *et al.*, 1999).

Na regressão logística utiliza-se frequentemente a técnica *stepwise*, a qual corresponde a um processo de entrada ou retirada de variáveis ao modelo, baseado em critérios como por exemplo o teste Wald. Neste caso, as variáveis com valores estatísticos (estatística Wald) mais elevados são os primeiros a ser incluídos no modelo.

No presente trabalho vai ser utilizado o método *Forward Stepwise* (Wald), através do qual, para obtenção do modelo final de regressão logística é necessário, no decorrer do processo, efectuar vários ajustes do modelo, ou seja, verificar se cada uma das variáveis obtidas nesse modelo está significativamente relacionada com a variável resposta do modelo. Esta análise é realizada através de testes de hipóteses estatísticos, que avaliam o modelo com e sem a variável.

Existem testes estatísticos para analisar a significância do modelo final, nomeadamente o teste Qui-quadrado e o teste Hosmer e Lemeshow.

Estas duas medidas, em combinação, fornecem suporte para que se aceite o modelo de regressão logística como significativo. São estes testes que asseguram a evidência de significância estatística das variáveis.

Primeiro realiza-se um teste Qui-quadrado para mudança no valor -2 Log Likelihood (-2 LL), do modelo base. Neste caso, a significância de um modelo é dada pela mudança verificada em LL (Log Likelihood), estatística de verosimilhança, a qual é multiplicada por -2 {-2LL}, de forma a assumir um número positivo e a ter uma distribuição aproximada ao Qui-quadrado (χ^2), antes e depois de se introduzirem as variáveis no modelo. A referida diferença é apresentada sob a forma de Qui-quadrado, podendo observar-se a significância estatística desse e assim concluir-se acerca da significância do modelo (Sharma, 1996).

Segundo Hair (1998), o ajuste geral do modelo pode ser avaliado utilizando-se a medida -2 LL. Se no modelo avaliado houver um decréscimo no valor -2LL comparado ao modelo base, existe melhoria no modelo, uma vez que menores valores da medida -2LL indicam o melhor ajuste do modelo.

Existem ainda três medidas de adequação do ajuste, as quais são comparáveis ao R^2 da regressão múltipla: A medida R^2 Cox e Snell, a medida Nagelkerke e medida Pseudo R^2 .

As medidas de adequação de ajuste comparam as probabilidades estimadas com as probabilidades observadas, para as quais valores mais altos indicam um ajuste melhor.

A medida R^2 Cox e Snell está limitada pelo facto de que não consegue alcançar o valor máximo de 1, de modo que Nagelkerke propôs uma modificação que tem o alcance de 0 para 1.

A medida Pseudo R^2 com base na melhoria do valor 2 LL.

O Pseudo R^2 é calculado como:

$$R^2_{\text{logn}} = \frac{2 \text{ LL base} - (-2 \text{ LL modelo})}{-2 \text{ LL base}}$$

Depois realiza-se a medida Hosmer e Lemeshow de ajuste geral, que tem um teste estatístico que indica se houve ou não diferença estatisticamente significativas entre as classificações observadas e previstas.

Mostra a não significância indicando a ausência de diferença na distribuição de valores dependentes efetivos e previstos.

Esta é a medida final do ajuste do modelo e mede a correspondência dos valores efetivos e previstos da variável dependente.

Neste caso, o melhor ajuste do modelo é indicado por uma diferença menor na classificação observada e prevista.

Também neste teste um bom ajuste de modelo é indicado por um valor Qui-quadrado não significante.

Caso a verdadeira distribuição de y condicionada por x , não seja logística, mas pertença à família exponencial, os estimadores obtidos com base no modelo “logit” continuarão a ser consistentes, o que permite alguma tranquilidade na utilização deste modelo.

4.4. Resultados

4.4.1. Análise Descritiva

Na realização do processo estatístico utilizou-se numa primeira fase uma Análise Univariada.

As medidas estatísticas analisadas foram a média, a mediana, o desvio-padrão, a kurtosis, o enviesamento e os casos extremos, o mínimo e o máximo.

Para que se possa perceber se se verificam diferentes comportamentos, ou comportamentos padrão nos rácios, ao nível dos dois subgrupos em estudo (clientes com risco e clientes sem risco), as medidas estatísticas associadas a cada rácio foram calculadas para cada subgrupo, atendendo aos intervalos de aplicabilidade dos modelos que serão estimados, na próxima secção.

As medidas estatísticas foram calculadas para todas as variáveis seleccionadas, com recurso ao software estatístico SPSS (Anexo 3). Após obtenção dos resultados, sentiu-se a necessidade de efetuar alguns reparos relativamente aos mesmos, uma vez que nem sempre se obtiveram resultados consistentes com o que seria de esperar para o subgrupo em análise. Deste modo, procurou-se seleccionar alguns rácios que se enquadrassem nesta situação, e pelo menos um rácio de cada tipologia.

Quadro 12 – Análise Descritiva - Rendibilidade do Ativo

Variável	Rendibilidade do Ativo (ROA)	
	Clientes Com Risco	Clientes Sem Risco
Média	-0,0196	0,0084
Mediana	0,0053	0,0103
Máximo	0,5830	3,4974
Mínimo	-2,4744	-2,4266
Desvio Padrão	0,1639	0,1813
Skewness	-5,7027	3,7515
Kurtosis	63,2494	177,9218

Como seria de esperar, a variável apresenta, em termos médios, um valor superior no grupo dos clientes sem risco, relativamente ao grupo oposto.

Aliás, neste último grupo, a média apresenta mesmo, valor negativo, o que indicia que as empresas possuem resultados negativos.

Também os valores extremos assumidos pela variável nos dois grupos, parecem delinear uma diferenciação entre os dois grupos, sendo que o valor máximo da rentabilidade do ativo nos clientes sem risco, é bem superior ao dos clientes com risco. A medida de skewness é negativa no grupo das empresas com risco e positiva no grupo oposto, o que sugere que no primeiro grupo as variáveis tendem a assumir valores à esquerda da média, por seu lado, no grupo de clientes sem risco, os valores observados para a variável, tendem predominantemente, a assumir valores à direita da média. Esta análise sobre o enviesamento, vem reforçar a hipótese de que a variável rentabilidade do ativo assume diferentes comportamentos, conforme se trate de um cliente com ou sem risco. A kurtosis sugere uma função de distribuição da variável leptocúrtica, o que significa que existem muitos valores que se afastam do valor médio.

Quadro 13 – Análise Descritiva - Rotação Ativo

Esta variável apresenta um comportamento muito semelhante, quer se tratem de clientes com risco, quer não. Por isso, não será em princípio, uma variável importante,

Variável Parâmetro	Rotação do Ativo	
	Clientes Com Risco	Clientes Sem Risco
Média	1,1143	1,1155
Mediana	0,7909	0,8996
Máximo	17,3076	18,4925
Mínimo	0,0000	0,0000
Desvio Padrão	1,3529	1,0823
Skewness	5,7814	5,6166
Kurtosis	55,0773	70,8432

na diferenciação de clientes com risco e sem risco. Espera-se que este rácio tenda a ver o seu valor reduzido, em empresas com algum risco, devido, nomeadamente, a uma queda das vendas.

Quadro 14 – Análise Descritiva - Aut. Financeira

Variável Parâmetro	Autonomia Financeira	
	Clientes Com Risco	Clientes Sem Risco
Média	1,0178	2,5394
Mediana	0,3211	0,6166
Máximo	142,2955	420,6056
Mínimo	-0,8668	-0,8992
Desvio Padrão	6,0347	16,5345
Skewness	19,6764	19,4464
Kurtosis	422,9553	444,7715

O capital próprio das empresas com risco tende a ser baixo ou mesmo negativo. O contrário tenderá a ocorrer ao nível do denominador deste rácio, ou seja, o

passivo total. Deste modo, é de esperar que a autonomia financeira seja em média maior nos clientes sem risco do que nos clientes com risco, conforme as médias calculadas nos dois grupos reflectem. Porém, verifica-se uma grande amplitude entre os extremos, e se a isso se juntar o desvio padrão elevado e a kurtosis leptocúrtica nos dois grupos, conclui-se que a diferença nas médias, por si só, não é estatisticamente significativa.

Quadro 15 - Análise Descritiva - Cob. Custos Financeir.

Variável	Cobertura dos Custos Financeiros	
	Clientes Com Risco	Clientes Sem Risco
Média	2.824,4124	-14,2458
Mediana	-0,7858	-0,2800
Máximo	2.776.203,0000	2.649,2268
Mínimo	-6.618,4167	-15.575,9079
Desvio Padrão	88.637,8101	555,5246
Skewness	31,3206	-23,1993
Kurtosis	980,9870	638,5809

Este rácio apresenta médias inconsistentes face às eventualmente esperadas. Seria normal que clientes com risco de entrar em incumprimento tivessem,

em média, uma cobertura dos custos financeiros mais reduzida do que os clientes sem risco, ou porque algumas empresas já não têm atividade e, portanto, os seus resultados operacionais encontram-se reduzidos ao mínimo ou são nulos, ou porque os empréstimos financeiros aumentaram. A análise detecta muita dispersão dos valores face às médias, não permitindo retirar qualquer conclusão consistente.

Por outro lado, é possível admitir que as empresas com risco, que possuam problemas financeiros graves, não contabilizem no curto prazo as dívidas de médio e longo prazo que se vão vencendo, como prevê o Plano Oficial de Contabilidade. Além disso, face às dificuldades financeiras, estas empresas tendem a não conseguir obter crédito bancário, seja de curto ou de médio/longo prazo.

Quadro 16 - Análise Descritiva - Liquidez Imediata

Variável	Liquidez Imediata	
	Clientes Com Risco	Clientes Sem Risco
Média	0,7014	0,9595
Mediana	0,0492	0,1180
Máximo	277,1525	90,1878
Mínimo	-0,0664	-0,6485
Desvio Padrão	9,0037	4,3201
Skewness	29,7045	13,6380
Kurtosis	909,9387	238,1793

O rácio de liquidez imediata tenderá a ser maior em empresas sem risco do que em empresas com risco, que deixem transparecer problemas financeiros, normalmente, empresas

com risco tendem a ter problemas de liquidez, apresentando tendencialmente, valores reduzidos de meios financeiros líquidos e dívida a terceiros elevada. Os valores extremos em conjunto com o desvio padrão e a kurtosis, evidenciam uma grande dispersão e afastamento dos valores observados da amostra face à média.

Quadro 17 - Análise Descritiva - Ativos Tangíveis

Os resultados obtidos para as medidas estatísticas não deixam transparecer qualquer diferenciação no comportamento da variável ao nível dos dois subgrupos em análise.

Variável	Ln (Ativos Tangíveis)	
	Clientes Com Risco	Clientes Sem Risco
Média	13,3239	13,2287
Mediana	13,4583	13,2859
Máximo	21,9727	22,7534
Mínimo	0,0000	0,0000
Desvio Padrão	2,4748	2,4010
Skewness	-1,2986	-1,1549
Kurtosis	6,9030	6,7950

Na realidade, esta variável não pode ser entendida solitariamente. Apesar de não demonstrar que a empresa possa vir a entrar em incumprimento, a variável pode ser utilizada como forma de prever se existem ativos que possam vir a ser executados, em caso de insolvência, podendo deste modo, vir a ser atenuado o nível de classificação de risco do cliente.

De um modo generalizado, os diferentes rácios apresentam médias diferentes em cada subgrupo, porém, as restantes medidas, nomeadamente o desvio padrão em cada um destes grupos, não assume valores, capazes de tornar cada uma daquelas diferenças, por si só, estatisticamente significativa.

Após esta primeira apreciação das variáveis, procurou-se identificar possíveis correlações entre elas, já que, por construção, alguns rácios não poderiam deixar de apresentar correlações significativas. Tratando-se de variáveis contínuas, pelo menos em termos teóricos, construiu-se a matriz dos coeficientes de correlação de Pearson (Anexo 4).

A análise desta matriz permite detectar um número significativo de pares de variáveis que apresentam elevados valores para os coeficientes de correlação respetivos, situação que terá algumas implicações na análise estatística posterior, nomeadamente em termos dos efeitos mais habituais da multicolinearidade, como sejam as variáveis cujo coeficiente aparece com sinal “trocado” ou a rejeição de variáveis teoricamente significativas.

Para Carvalho das Neves et al. (1998), se as consequências positivas de não se eliminar “à priori” algumas variáveis são fáceis de entender já que assim se poderá ter uma maior latitude na fase de modelização, é também importante sublinhar que esta opção também implica algumas consequências menos boas, sobretudo, tendo em conta as metodologias habitualmente utilizadas neste tipo de análise. De facto, a generalidade dos modelos é construída com base em metodologias do tipo “stepwise”, isto é, em que a escolha dos rácios financeiros a incluir nos modelos apenas se fundamenta num maior grau de ajustamento, sem um sólido suporte de uma modelização teórica. Neste contexto, o manter na base de dados variáveis com elevadas correlações desaconselha, de todo o recurso a metodologias de “stepwise” de tipo “backward” (inclusão de todas as variáveis num primeiro tempo e eliminação, passo a passo, daquelas que se venham a mostrar sem significância estatística) e mesmo no recurso a uma metodologia “forward” ou a um modelo de entrada direta pode arrastar problemas de multicolinearidade que se poderão manifestar, nomeadamente, nos sinais dos coeficientes. Para além deste último problema, também se deve assinalar que a multicolinearidade também pode induzir alguma

instabilidade nas estimativas obtidas para os diferentes coeficientes, particularmente se se utilizar um “stepwise”.

Tendo no entanto presente o estado de conhecimentos teóricos sobre este tipo de modelos, a solução de manter, pelo menos numa primeira fase, todas as variáveis pareceu constituir a solução mais adequada.

4.4.2. Modelo de Regressão Logística - Logit

Após a análise Univariada das variáveis realizou-se uma análise Multivariada, onde foi utilizado, como anteriormente se referiu, um modelo de escolha binária – Modelo de Regressão Logística (*Logit*).

As tabelas seguintes descrevem o *output* do programa SPSS na análise estatística do modelo de regressão *logit*, - no qual se usou o método stepwise forward (wald), para definir o modelo final. Este método tem por objetivo minimizar o número de variáveis e maximizar a precisão do modelo.

Conforme referido na secção anterior, vão ser utilizadas no processo de modelação as 31 variáveis mencionadas, através das quais se vai tentar, conforme objetivo deste trabalho, determinar as variáveis quantitativas que melhor consigam prever as situações de futuros incumprimentos de clientes na empresa comercializadora de energia.

Um ponto importante na realização da regressão logística é a definição do ponto de corte. Quando a amostra reflete a proporção existente na população, o ponto de corte que se utiliza corresponde à taxa verificada entre as duas situações em análise (clientes com risco e sem risco).

Como neste caso, não existe uma efetiva certeza sobre a proporcionalidade dos dois casos em estudo, face à população, porque o critério utilizado na selecção da amostra privilegiou antes a proporcionalidade dos segmentos de negócio, irá utilizar-se o ponto de corte de 0,5 que define probabilidades iguais para as duas situações em análise.

Recordando o que foi sustentado na subsecção 4.1 sobre a segmentação da amostra, considerou-se que a mesma poderia encontrar-se no VN de um milhão de euros, ou de dois milhões, deste modo, as duas opções foram testadas, tendo-se mesmo, para efeitos de reforço de sustentação da decisão tomada, realizado sucessivas modelações em SPSS, atendendo a diferentes segmentações, ou seja, utilizando diferentes intervalos de

volume de negócios. Assim, foram realizadas estimações atendendo aos seguintes critérios de segmentação.

- Modelação Sem Segmentação;
- Modelação Com Segmentação - 1 Milhão de Euros VN;
- Modelação Com Segmentação - 2 Milhões de Euros VN;
- Modelação Com Segmentação - 10 Milhões de Euros VN;
- Modelação Com Segmentação - 50 Milhões de Euros VN.

A escolha final dos modelos, acabou por recair sobre a segmentação de um milhão de euros, por ser aquela, onde a precisão do modelo é maior face às restantes. De facto, de entre todas as opções testadas, as que consideraram como referência, volume de negócios mais baixos (um e dois milhões de euros), foram aquelas onde se obtiveram maiores taxas de acerto, ou seja, maior precisão. Na realidade, a opção pelo modelo com “zona de corte” da segmentação no um milhão de euros de VN, preterindo o de dois milhões, deveu-se essencialmente ao facto do primeiro apresentar uma distribuição mais equilibrada das taxas de acerto ao nível dos dois subgrupos, e também, por possuir uma taxa de acerto dos casos de incumprimento superior. As restantes opções, onde foram considerados níveis de VN superiores como referência da segmentação, revelaram modelos menos ajustados e precisos.

De acordo com o exposto, obtiveram-se dois modelos distintos, devendo um aplicar-se a empresas com um volume de negócios anual inferior a um milhão de euros (**Modelo I**), e o outro, a empresas com um volume superior ao referido (**Modelo II**).

Através do output do SPSS, serão apresentados os dois modelos, constando no Anexo 5, uma síntese dos resultados obtidos em outras modelações excluídas, pelos motivos já referidos.

4.4.2.1. Modelo I - Clientes Com Vol. Negócios Inferior a Um Milhão de Euros

A amostra contém 767 observações de clientes que possuem volume de negócios anual inferior a 1 milhão de euros.

❑ Introdução de Variáveis

Nos passos seguintes serão introduzidas sequencialmente variáveis no modelo de forma a se obter, o melhor modelo. A primeira variável a ser introduzida no modelo será aquela que tiver a estatística de pontuação mais alta, estatística de Wald.

As variáveis foram selecionadas em 6 etapas, ou seja, a estatística Wald utilizou 6 passos até obter o modelo final.

Através das significâncias estatísticas do modelo (*estatística do Qui-Quadrado*), constata-se que os coeficientes são significativos em cada passo. Logo, o modelo é estatisticamente significativo, rejeitando-se desta forma a hipótese nula de que as variáveis independentes não são explicativas da variável dependente Y (Clientes Com Risco – Em incumprimento).

❑ Avaliação do Ajuste Geral do Modelo

O Quadro 18 apresenta três medidas para analisar o ajuste geral do modelo:

- - 2 Log Likelihood
- Cox & Snell (a qual está limitada visto que não obtém o valor 1)
- Nagelkerke (corrige a anterior e já apresenta um domínio de 0 a 1)

Quadro 18 – Medidas de Ajustamento do Modelo

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1048,979 ^a	,014	,019
2	1040,027 ^a	,026	,035
3	1032,388 ^a	,036	,047
4	1027,036 ^b	,042	,056
5	1022,916 ^b	,047	,063
6	1016,251 ^b	,056	,074

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than ,001.

b. Estimation terminated at iteration number 5 because parameter estimates changed by less than ,001.

Pela análise do quadro, verifica-se que, à medida que foram introduzidas variáveis no modelo, houve uma redução no valor de -2 Log Likelihood, indicando uma melhoria no modelo.

O valor diminuiu de 1048,979 no passo 1 para 1016,251 no passo 6.

Esta diminuição determina que o modelo apresentado se encontra ajustado e o valor de 1016,251 apresentado no último passo (passo 6) corresponde à medida geral em como o modelo se ajusta.

De acordo com Hair (2005), um modelo bem ajustado possui um valor pequeno para - 2 LL, sendo o seu valor mínimo zero.

Contrariamente à medida - 2 Log Likelihood, no caso do R^2 de Cox & Snell e do R^2 de Nagelkerke, a melhoria do modelo é traduzida pelo aumento do valor destas medidas à medida que se introduzem variáveis no modelo.

O valor R^2 de Cox & Snell varia de 0,014 para 0,056 e o valor R^2 de Nagelkerke (que corrige a anterior para poder apresentar um domínio de 0 a 1), varia de 0,019 para 0,074. Como se referiu anteriormente o aumento do valor destas medidas traduz uma melhoria no poder explicativo do modelo.

Segundo Hair (2005), estas duas medidas comparam as probabilidades estimadas com as probabilidades observadas, sendo que valores mais altos significam um melhor ajuste do modelo.

Assim, verifica-se pelo quadro, que à medida que foram introduzidas variáveis no modelo, houve um aumento do valor nestas duas medidas, ou seja, à medida que foram introduzidas variáveis ao modelo, este foi-se ajustando melhor, de forma a apresentar valores mais altos.

Como foi referido anteriormente, a última medida de ajuste do modelo é o valor de Hosmer and Lemeshow Test, (Quadro 19) o qual mede a correspondência entre valores reais e os previstos da variável dependente. Possui um modelo estatístico que indica se ocorreram diferenças estatisticamente significativas entre as classificações observadas e previstas.

Nesta medida, o melhor ajuste do modelo é indicado por uma diferença menor na classificação observada e prevista. Um bom ajuste do modelo é dado por um valor de Qui-quadrado não significante.

Quadro 19 – Teste Hosmer and Lemeshow

Step	Chi-square	df	Sig.
1	12,373	8	,135
2	15,127	8	,057
3	14,265	8	,075
4	21,214	8	,007
5	11,139	8	,194
6	15,663	8	,047

Estas medidas combinadas (-2LL e Hosmer and Lemeshow Test), sugerem a aceitação do modelo do último passo como um modelo signficante de regressão logística.

Pelos resultados apresentados podemos concluir que o modelo obtido no passo seis se encontra ajustado. Aceita-se este modelo de Regressão Logístico, como significativo.

❑ Avaliação da Precisão Estimativa do Modelo

Após se proceder à avaliação do Ajuste Geral do Modelo, segue-se uma avaliação da Precisão da Estimativa do Modelo, para a qual é utilizado o Método das Matrizes de Classificação.

O Quadro 20 constitui a matriz de classificação, a qual serve para avaliar a precisão de previsão do modelo.

Pela análise do quadro (matriz de classificação), verifica-se a classificação dos casos preditos pelo modelo, comparando com os dados reais inerentes à variável dependente Y.

Quadro 20 – Tabela de Classificação^a

Observed			Predicted		
			Y		Percentage Correct
			0	1	
<u>Step 1</u>	Y	0	48	311	13,4
		1	36	372	91,2
		Overall Percentage			54,8
<u>Step 2</u>	Y	0	158	201	44,0
		1	122	286	70,1
		Overall Percentage			57,9
<u>Step 3</u>	Y	0	153	206	42,6
		1	108	300	73,5
		Overall Percentage			59,1
<u>Step 4</u>	Y	0	176	183	49,0
		1	114	294	72,1
		Overall Percentage			61,3
<u>Step 5</u>	Y	0	181	178	50,4
		1	115	293	71,8
		Overall Percentage			61,8
<u>Step 6</u>	Y	0	153	206	42,6
		1	92	316	77,5
		Overall Percentage			61,1

a. The cut value is ,500

O modelo apresenta uma melhor explicação no caso de incumprimento do cliente, do que nas situações onde aparentemente o cliente cumpre. São explicados corretamente 42,6% dos casos de clientes que não incumprem e 77,5% dos incumprimentos, ou seja, o modelo explica corretamente o incumprimento de 316 clientes, num total de 408 casos de incumprimento, e acerta em 206 casos de cumprimento, num total de 359 clientes cumpridores.

Esta menor explicação no caso dos clientes que cumprem, pode estar associada ao facto de terem sido classificados como clientes “Sem Risco”, empresas que eventualmente até já apresentam alguma debilidade financeira, no entanto, mantém os pagamentos das faturas de energia em dia. Tratando-se de um fornecimento chave para o prosseguimento do negócio, ainda que com dificuldades, normalmente, o fornecedor de energia é um dos últimos que deixa de ser pago. Estas situações ocorrem sobretudo em casos que geralmente, degeneram em insolvência ou até mesmo, falência das empresas.

Por outro lado, convém referir que apesar de satisfatória a taxa de acerto (77,5%), no caso dos clientes considerados “Com Risco”, acredita-se que a mesma poderia ter sido mais elevada. Também aqui, existem algumas ressalvas relativamente aos critérios utilizados para classificar os clientes “Com Risco” ou em incumprimento. Recorde-se que foram

considerados clientes “Com Risco”, todos aqueles que possuísem registo em contencioso, e que apresentassem simultaneamente, dívida vencida e ausência de registos de reclamações ou pedidos, e este último critério poderá ser falível em alguns casos. Verificam-se essencialmente dois motivos para que isso possa ocorrer: Primeiro, a base de dados de reclamações e pedidos nem sempre se encontra devidamente atualizada, pelo que o cliente pode ter faturas em atraso, sobretudo nos casos de dívida vencida à menos de 6 meses, mas por motivos que lhe são alheios, e segundo, registam-se alguns casos de incumprimento “intencional”, em que o cliente deixa vencer o pagamento de uma ou várias faturas sistematicamente, acabando por liquidar o total da sua dívida posteriormente. De um modo geral, este fenómeno ocorre com grandes clientes.

Para avaliar melhor este tipo de situações, seria necessário possuir um historial de comportamento de pagamentos do cliente, esta informação ajudaria a desmistificar aquele tipo de situações.

Estes casos concorrem obviamente, para a taxa de acerto do modelo, que decerto teria sido superior, se as situações descritas tivessem sido minuciosamente deslindadas, no momento da classificação dos clientes.

Na aplicação do modelo deve-se ainda salientar o aspecto da amostra não possuir o mesmo número de observações para clientes “Com Risco” e para clientes “Sem Risco”, principalmente o facto do número de clientes “Com Risco” da amostra não representar a mesma proporção que a população, sendo que, por norma, o número de incumprimentos, é bastante inferior à situação oposta.

No entanto, considerando todo o modelo, verifica-se que a taxa de acerto geral é de 61,1%, ou seja, mais de metade das situações estão corretamente explicadas num dos dois grupos do modelo.

Conclui-se que, no geral, o modelo apresenta um poder explicativo razoável, embora seja mais explicativo no caso das situações de incumprimento. Convém aqui relembrar, que este é um aspecto muito positivo, uma vez que, como foi evidenciado no capítulo 3, os erros de tipo I (classificar um cliente “com risco”, como “sem risco”), acarretam sempre custos mais elevados do que os erros de tipo II (classificar um cliente “sem risco” como “com risco”).

Realça-se ainda, que para casos em que as amostras em análise não possuem o mesmo tamanho (como é o caso, em que a amostra não possui o mesmo número de observações

para clientes com risco e sem risco), é conveniente identificar qual a percentagem mínima de classificação que é aceitável para o modelo.

Hair (2005) define um critério de proporcionalidade, o qual tem justamente em consideração a diferença de tamanho das amostras, com o qual vai identificar a percentagem mínima que é aceitável para o modelo.

Neste modelo possuímos 359 observações de clientes “Sem Risco” ($Y=0$) e 408 observações de clientes “Com Risco” ($Y=1$).

A fórmula do critério proporcional é a seguinte:

$$CP = p^2 + (1 - p)^2$$

CP – Critério proporcional

p – proporção de clientes “Sem Risco”

1 – p = proporção de clientes “Com Risco”

Aplicando a fórmula obtemos:

$$CP = (0,468)^2 + (0,532)^2$$

$$CP = 0,502$$

Como se observou anteriormente, a precisão de classificação do modelo é 61,1%, superior ao critério proporcional obtido, daí considera-se como aceitável o modelo estimado.

❑ Validação do Modelo

De acordo com o referido na revisão de literatura, a validação do modelo de regressão logística pode ser obtida através da criação de amostras de treino e de validação (Hair *et al.*, 1998).

Como amostra de validação utilizou-se uma amostra obtida aleatoriamente pelo programa SPSS, para a qual se utilizou 200 operações (n.º escolhido a título de exemplo). A validação do modelo de regressão logística é obtida através da aplicação do modelo na amostra de validação.

Os quadros seguintes mostram a tabela de classificação com os resultados obtidos na amostra de treino (Quadro 21), os quais se poderão comparar com os resultados já apresentados no quadro 22, da amostra de validação.

Quadro 21 – Amostra de Treino – Matriz de Classificação

		Predicted		
		Y		Percentage Correct
		0	1	
Step 6	Observed	0	1	
	Y	153	206	42,6
		92	316	77,5
	Overall Percentage			61,1

Quadro 22 – Amostra de Validação – Matriz de Classificação

Observed		Predicted			
		Y		Percentage Correct	
		0	1		
Step 6	Y	0	55	26	67,90
		1	48	71	59,66
Overall Percentage					63,00

No cômputo geral do modelo, verifica-se que a taxa de acerto é semelhante, apresentando-se até um pouco mais elevada na amostra de validação. Porém, verifica-se que na amostra de validação o modelo prevê melhor o não incumprimento (67,9%) do que o próprio incumprimento, ao contrário do que ocorre na amostra de treino (42,6%).

Verificando os dados gerais das duas amostras pode concluir-se que o modelo de regressão logística possui um forte suporte empírico tanto na amostra de validação como na de treino para explicar as variáveis dependentes obtidas.

❑ Variáveis na Equação

O Quadro 23 representa as variáveis finais obtidas pela regressão logística, e os respectivos Betas, ou seja apresenta os vários parâmetros presentes na equação de regressão logística.

É possível verificar ainda, o valor da estatística *Wald*, que identifica o quanto cada variável independente participa individualmente para a explicação da variável dependente. Observa-se pelo teste *Wald* que a significância de todos os parâmetros foi aceitável. O modelo final selecionou seis variáveis das trinta e duas inicialmente incluídas.

Quadro 23 – Variáveis do Modelo

		B	S.E.	Wald	Df
Step 1^a	Vendas	-,083	,026	9,921	1
	Constant	1,129	,329	11,777	1
Step 2^b	Vendas	-,079	,026	8,936	1
	Rácio do FM	-,401	,145	7,641	1
	Constant	1,106	,329	11,336	1
Step 3^c	Solvabilidade	-,016	,007	4,714	1
	Vendas	-,093	,028	11,010	1
	Rácio do FM	-,363	,144	6,321	1
	Constant	1,335	,355	14,135	1
Step 4^d	RemunCapitaisAlheios	5,785	2,571	5,063	1
	Solvabilidade	-,017	,008	4,798	1
	Vendas	-,101	,029	12,386	1
	Rácio do FM	-,394	,145	7,335	1
	Constant	1,267	,361	12,317	1
Step 5^e	RemunCapitaisAlheios	5,827	2,553	5,208	1
	Solvabilidade	-,015	,007	4,044	1
	PesoEndLP	,478	,236	4,094	1
	Vendas	-,096	,029	11,030	1
	Rácio do FM	-,455	,151	9,053	1
	Constant	1,068	,373	8,190	1
Step 6^f	RemunCapitaisAlheios	5,859	2,581	5,152	1
	Solvabilidade	-,012	,007	2,739	1
	PesoEndLP	1,817	,584	9,684	1
	Vendas	-,094	,029	10,448	1
	Rácio do FM	-,333	,151	4,893	1
	EstruturadoEndividamento	1,534	,607	6,381	1
	Constant	-,315	,664	,224	1

a. Variable(s) entered on step 1: Vendas.

b. Variable(s) entered on step 2: Rácio do FM.

c. Variable(s) entered on step 3: Garantia.

d. Variable(s) entered on step 4: RemunCapitaisAlheios.

e. Variable(s) entered on step 5: PesoEndLP.

f. Variable(s) entered on step 6: EstruturadoEndividamento.

Os resultados demonstram que existe uma adequação do modelo para explicar o comportamento da variável dependente, onde as variáveis que resultaram no modelo final: Remuneração dos Capitais Alheios, rácio de Solvabilidade, Peso do Endividamento de Longo Prazo, Vendas, Rácio de Fundo de Maneio e Estrutura do Endividamento, são as que melhor explicam a variável dependente, ou seja, a probabilidade de incumprimento.

□ Aplicação da equação do Modelo e Cálculo da Probabilidade

Assim, a partir do resultado dos parâmetros estimados pelo modelo *logit*, obtemos a seguinte equação final.

$$Y_i^* = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i = X_i \beta + u_i$$

Onde, Y vem em função de **Y*** :

$$Y_i = \begin{cases} 1, & \text{Se } Y^* \geq 0,5 \\ 0, & \text{Se } Y^* < 0,5 \end{cases}$$

$$Y^* = -0,315 + 5,859 \times \text{Remun. Cap. Alheios} - 0,12 \times \text{Solvabilidade} + 1,817 \times \text{Peso End. LP} - 0,094 \times \text{Vendas} - 0,333 \times \text{Rácio Fundo Maneio} + 1,534 \times \text{Est. Endividamento}$$

Aplicando os respetivos valores das variáveis no modelo de regressão linear, obtém-se para cada variável um valor de probabilidade de ocorrência de incumprimento, o qual como vimos anteriormente é definido por:

$$\text{Prob}(Y_i = 1) = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki})}}$$

O Quadro 24 apresenta para vários intervalos, as probabilidades obtidas de ocorrência de incumprimento. Relembre-se que no desenvolvimento do modelo se utilizou um ponto de corte de 0,5.

Quadro 24 – Probabilidades de Incumprimento

Intervalo Probabilidade	Cientes "Sem Risco"	Cientes "Com Risco"	Total
0,00 a 0,25	58		58
0,25 a 0,50	356		356
0,50 a 0,75		315	315
0,75 a 1,00		38	38
Total	414	353	767

Pela análise do quadro é possível verificar que existe entre 50 a 75% de probabilidade de 315 clientes entrarem em incumprimento e entre 75 e 100% de entrarem 38.

4.4.2.2. Modelo II – Clientes Com Vol. Negócios Superior a Um Milhão de Euros

A amostra contém 1195 observações de clientes que registam um valor de negócios anual superior a 1 milhão de euros.

❑ Introdução de Variáveis

Nos passos seguintes serão introduzidas sequencialmente variáveis no modelo de forma a se obter o melhor modelo. A primeira variável a ser introduzida no modelo será aquela que tiver a estatística de pontuação mais alta, estatística de wald.

As variáveis foram selecionadas em 6 etapas, ou seja, a estatística Wald utilizou 6 passos até obter o modelo final.

Através das significâncias estatísticas do modelo (*estatística do Qui-Quadrado*), constata-se que os coeficientes são significativos em cada passo. Logo, o modelo é estatisticamente significativo, rejeitando-se desta forma a hipótese nula de que as variáveis independentes não são explicativas da variável dependente Y (Clientes Com Risco – Em incumprimento).

❑ Avaliação do Ajuste Geral do Modelo

O Quadro 25 apresenta três medidas para analisar o ajuste geral do modelo:

- - 2 Log Likelihood
- Cox & Snell (a qual está limitada visto que não obtém o valor 1)
- Nagelkerke (corrige a anterior e já apresenta um domínio de 0 a 1)

Quadro 25 – Medidas de Ajustamento do Modelo

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1596,574 ^a	,046	,062
2	1573,725 ^b	,064	,086
3	1569,405 ^b	,068	,091
4	1564,740 ^a	,071	,095
5	1567,072 ^a	,070	,093
6	1561,446 ^a	,074	,099

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than ,001.

b. Estimation terminated at iteration number 5 because parameter estimates changed by less than ,001.

Pela análise do quadro, verifica-se que, à medida que foram introduzidas variáveis no modelo, houve uma redução no valor de -2 Log Likelihood, indicando uma melhoria no

modelo. Esta diminuição determina que o modelo apresentado se encontra ajustado e o valor de 1561,446 apresentado no último passo (passo 6) corresponde à medida geral em como o modelo se ajusta.

Nas medidas do R^2 de Cox & Snell e do R^2 de Nagelkerke, uma melhoria do modelo traduz-se pelo aumento do valor destas medidas, à medida que se introduzem variáveis no modelo. O aumento do valor destas medidas traduz uma melhoria no poder explicativo do modelo.

A última medida de ajuste do modelo que será referida é o valor de Hosmer and LemeshowTest, no Quadro 26.

Relembre-se que, nesta medida, o melhor ajuste do modelo é indicado por uma diferença menor na classificação observada e prevista. Um bom ajuste do modelo é dado por um valor de Qui-quadrado não significativo.

Quadro 26 – Teste Hosmer and Lemeshow

Step	Chi-square	df	Sig.
1	5,759	8	,674
2	11,858	8	,158
3	18,660	8	,017
4	21,377	8	,006
5	24,878	8	,002
6	16,366	8	,037

Estas medidas combinadas (-2LL e Hosmer and Lemeshow Test), sugerem a aceitação do passo 6 como um modelo significativo de regressão logística.

Pelos resultados apresentados conclui-se que o modelo obtido no passo 6 se encontra ajustado. Aceita-se este modelo de Regressão Logístico, como significativo.

❑ Avaliação da Precisão Estimativa do Modelo

O Quadro 27 constitui a matriz de classificação, a qual serve para avaliar a precisão de previsão do modelo.

Pela análise do quadro (matriz de classificação), verifica-se a classificação dos casos preditos pelo modelo, comparando com os dados reais inerentes à variável dependente Y.

Quadro 27 – Tabela de Classificação^a

Observed			Predicted		
			Y		Percentage Correct
			0	1	
Step 1	Y	0	425	196	68,4
		1	291	282	49,2
		Overall Percentage			59,2
Step 2	Y	0	419	202	67,5
		1	248	325	56,7
		Overall Percentage			62,3
Step 3	Y	0	421	200	67,8
		1	236	337	58,8
		Overall Percentage			63,5
Step 4	Y	0	403	218	64,9
		1	217	356	62,1
		Overall Percentage			63,6
Step 5	Y	0	391	230	63,0
		1	214	359	62,7
		Overall Percentage			62,8
Step 6	Y	0	395	226	63,6
		1	224	349	60,9
		Overall Percentage			62,3

a. The cut value is ,500

Ao contrário do modelo anterior, este apresenta uma ligeira melhor explicação no caso dos clientes que cumprem, do que nas situações onde se verifica incumprimento. O modelo explica corretamente 63,6% dos casos de clientes que não incumprem e 60,9% dos incumprimentos, ou seja, explica corretamente as situações de 395 clientes que cumprem, num total de 621 casos de cumprimento das obrigações, e acerta em 349 casos de incumprimento, num total de 573 clientes que entraram em incumprimento.

Convém referir, neste momento, que um dos principais motivos pelo qual foi rejeitado o modelo com “zona de corte” nos 2 milhões de euros foi o facto do modelo aplicável a empresas com VN superior a 2 milhões de euros, possuir uma taxa de acerto de incumprimentos muito baixa (33%). Embora a sua taxa de acerto nos casos normais de cumprimento fosse muito boa (84%), a previsão de incumprimentos sairia negligenciada. Com o modelo selecionado, pode-se ter perdido algum poder de previsão de não incumprimentos, no entanto, essa diminuição é compensada pelo aumento de previsão dos casos de incumprimento.

Mais uma vez, valerá a pena relembrar dois motivos essenciais, que poderão ter contribuído para que as taxas de acerto não tenham sido superiores. Ao nível da taxa de acerto dos incumprimentos: Primeiro, a base de dados de reclamações e pedidos nem

sempre se encontra devidamente atualizada, pelo que o cliente pode ter faturas em atraso, sobretudo nos casos de dívida vencida à menos de 6 meses, mas por motivos que lhe são alheios, e portanto, nesses casos o incumprimento não lhes pode ser imputado, e segundo, registam-se alguns casos de incumprimento “intencional”, em que o cliente deixa vencer o pagamento de uma ou várias faturas sistematicamente, acabando por liquidar o total da sua dívida posteriormente. Como também já se referiu, de um modo geral, este fenómeno ocorre com grandes clientes. Se se pensar que este modelo é aplicável a clientes com volume de negócios superiores a 1 milhão de euros, e portanto, contempla todos os grandes clientes, é muito provável que a amostra contenha algumas situações semelhantes à descrita.

Estes casos concorrem obviamente, para a taxa de acerto do modelo, que decerto teria sido superior, se as situações descritas tivessem sido identificadas no momento da classificação dos clientes, no entanto, o facto de não existir um estudo analítico ou uma base de dados credível, que permita acompanhar a assiduidade e pontualidade do comportamento de pagamentos do cliente, impediu a identificação dessas situações e constitui uma limitação do trabalho.

Na aplicação do modelo deve-se ainda salientar o aspecto da amostra não possuir o mesmo número de observações para clientes “Com Risco” e para clientes “Sem Risco”, principalmente o facto do número de clientes “Com Risco” da amostra não representar a mesma proporção que a população, sendo que, por norma, o número de incumprimentos, é bastante inferior à situação oposta.

Avaliando o modelo no seu todo, verifica-se que a taxa de acerto geral é de 62,3%, ou seja, mais de metade das situações estão correctamente explicadas num dos dois grupos do modelo.

Conclui-se que no geral, o modelo apresenta um poder explicativo razoável, embora seja mais explicativo no caso das situações de não incumprimento.

Neste modelo, ao contrário do anterior, o erro de tipo I (classificar um cliente “com risco” como “sem risco”) é superior ao erro tipo II (classificar um cliente “sem risco” como “com risco”), o que constitui um aspecto menos bom, conforme já foi evidenciado em momentos anteriores deste trabalho.

À semelhança do Modelo I, a amostra não possui o mesmo número de observações para clientes com risco e sem risco, tendo-se calculado a percentagem mínima de classificação que é aceitável para o modelo.

Neste modelo tem-se 621 observações de clientes “Sem Risco” (Y=0) e 573 observações de clientes “Com Risco” (Y=1).

A fórmula do critério proporcional é a seguinte:

$$CP = p^2 + (1 - p)^2$$

CP – Critério proporcional

p – proporção de clientes “Sem Risco”

1 – p = proporção de clientes “Com Risco”

Aplicando a fórmula obtemos:

$$CP = 0,501$$

Como se observou anteriormente, a precisão de classificação do modelo é 62,3%, superior ao critério proporcional obtido, daí considera-se como aceitável o modelo estimado.

❑ Validação do Modelo

No processo de validação deste modelo, aplicou-se os mesmos procedimentos, do que no anterior, tendo-se utilizado uma amostra obtida aleatoriamente pelo programa SPSS, para a qual se utilizou 200 clientes (n.º escolhido a título de exemplo). A validação do modelo de regressão logística é obtida através da aplicação do modelo na amostra de validação.

Os quadros seguintes mostram a tabela de classificação com os resultados obtidos na amostra de validação (Quadro 29), os quais se poderão comparar com os resultados apresentados no quadro 28, da amostra de treino.

Quadro 28 – Amostra de Treino – Matriz de Classificação

Observed		Predicted		
		Y		Percentage Correct
		0	1	
Step 6	Y	395	226	63,6
		224	349	60,9
Overall Percentage				62,3

Quadro 29 – Amostra de Validação – Matriz de Classificação

Observed		Predicted		
		Y		Percentage Correct
		0	1	
Step 6	Y	68	30	69,4
		36	66	64,7
Overall Percentage				67,0

No geral pode-se verificar que as taxas de acerto na amostra de validação são semelhantes às taxas de acerto na amostra de treino.

Analisando os dados gerais das duas amostras, pode-se concluir que o modelo de regressão logística possui um forte suporte empírico tanto na amostra de validação como na de treino para explicar as variáveis dependentes obtidas.

❑ Variáveis na Equação

O Quadro 30 representa as variáveis finais obtidas pela regressão logística, e os respectivos Betas, ou seja apresenta os vários parâmetros presentes na equação de regressão logística.

Pode-se ver ainda o valor da estatística *Wald*, que identifica o quanto cada variável independente participa individualmente para a explicação da variável dependente. Observa-se pelo teste *Wald* que a significância de todos os parâmetros foi aceitável. Como é possível verificar o modelo final selecionou quatro variáveis das trinta e duas inicialmente incluídas.

Quadro 30 – Variáveis do Modelo

		B	S.E.	Wald	df
Step 1 ^a	RácioloFM	-1,417	,201	49,636	1
	Constant	,095	,065	2,179	1
Step 2 ^b	RácioloFM	-,660	,258	6,535	1
	Endividamento	1,423	,311	20,981	1
	Constant	-,957	,237	16,259	1
Step 3 ^c	PesoEndLP	,513	,247	4,311	1
	RácioloFM	-,773	,263	8,601	1
	Endividamento	1,242	,321	14,990	1
	Constant	-,953	,237	16,229	1
Step 4 ^d	PesoEndLP	,568	,249	5,199	1
	LiquidezGeral	-,096	,050	3,661	1
	RácioloFM	-,466	,304	2,354	1
	Endividamento	1,118	,323	12,003	1
	Constant	-,754	,255	8,731	1
Step 5 ^e	PesoEndLP	,517	,247	4,383	1
	LiquidezGeral	-,137	,047	8,569	1
	Endividamento	1,340	,288	21,613	1
	Constant	-,878	,246	12,760	1
Step 6 ^e	PesoEndLP	,516	,247	4,370	1
	LiquidezGeral	-,126	,046	7,423	1
	Endividamento	1,346	,286	22,095	1
	Cober. AtivoFixoTangível	,072	,035	4,187	1
	Constant	-,952	,247	14,924	1

- a. Variable(s) entered on step 1: RácioloFM.
b. Variable(s) entered on step 2: Endividamento.
c. Variable(s) entered on step 3: PesoEndLP.
d. Variable(s) entered on step 4: LiquidezGeral.
e. Variable(s) entered on step 6: CoberturaAtivoFixoTangível.

Os resultados demonstram que existe uma adequação do modelo para explicar o comportamento da variável dependente, onde as variáveis que resultaram no modelo final: Endividamento, Peso do Endiv. LP, Liquidez Geral e Cobertura do Ativo Tangível, são as que melhor explicam a variável dependente.

□ Aplicação da Equação do Modelo e Cálculo da Probabilidade

Assim, a partir do resultado dos parâmetros estimados pelo modelo *logit*, obtém-se a seguinte equação final :

$$Y_i^* = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i = X_i \beta + u_i$$

Onde, Y vem em função de Y* :

$$Y_i = \begin{cases} 1, & \text{Se } Y^* \geq 0,5 \\ 0, & \text{Se } Y^* < 0,5 \end{cases}$$

$$Y^* = -0,952 + 0,516 \times \text{Peso Endiv. LP} - 0,126 \times \text{Liquidez Geral} + 1,346 \times \text{Endividamento} + 0,072 \times \text{Cob. Ativo Fixo Tangível}$$

Aplicando os respectivos valores das variáveis no modelo de regressão linear, obtém-se para cada variável um valor de probabilidade de ocorrência de incumprimento, o qual como vimos anteriormente é definido por:

$$\text{Prob}(Y_i = 1) = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki})}}$$

O Quadro 31 apresenta para vários intervalos, as probabilidades obtidas de ocorrência de incumprimento. Devemos ter em atenção que no desenvolvimento do modelo utilizou-se um ponto de corte de 0,5.

Quadro 31 – Probabilidades de Incumprimento

Intervalo Probabilidade	Clientes "Sem Risco"	Clientes "Com Risco"	Total
0,00 a 0,25	68		68
0,25 a 0,50	549		549
0,50 a 0,75		551	551
0,75 a 1,00		26	26
Total	617	577	1194

Pela análise do quadro é possível verificar que existe entre 50 a 75% de probabilidade de 551 clientes entrarem em incumprimento e entre 75 e 100% de entrarem 26.

Capítulo 5 - Conclusões Finais

5.1. Principais Conclusões

O presente trabalho foi elaborado com a perspetiva de produzir um conjunto de inputs, que permitam à empresa encetar uma política de avaliação do risco de crédito dos seus clientes.

No sentido de aproveitar sinergias, procurou-se numa primeira fase, testar a aplicabilidade dos dois modelos de avaliação do risco de crédito utilizados noutra empresa do grupo.

Os resultados obtidos não se revelaram satisfatórios, uma vez que nenhum dos modelos conseguiu prever as situações de clientes cumpridores, tendo ambos os modelos, classificado todos os clientes da amostra como “Com Risco”.

Estes resultados atribuem-se ao facto dos modelos testados terem sido criados para uma carteira de clientes com características diferentes. De facto, os modelos aplicam-se a empresas que se inserem num tecido empresarial com uma dimensão maior e uma estrutura setorial mais ampla e diversificada.

Para além disso, verificam-se constrangimentos de ordem contabilística, podendo as rubricas que compõem os rácios não possuir a mesma interpretação nas duas geografias a que as empresas pertencem.

Refira-se ainda, o facto da informação sobre o volume de negócios de todas as empresas clientes não se encontrar disponível antes da seleção da amostra. No caso de se possuir esta informação, a amostra de clientes a aplicar nos testes aos modelos iniciais poderia ter sido mais direcionada e ampla. Julga-se por isso necessário, que em futuros trabalhos de âmbito semelhante, se possua aquela informação para toda a carteira de clientes, aliás, não se pode deixar de referir, que se trata de informação relevante para o próprio negócio, podendo o gestor delinear a sua estratégia de abordagem ao cliente, com base nessa mesma informação.

Na sequência dos critérios inerentes aos modelos (rácios financeiros e respetivas ponderações, limites de volume de negócios para aplicabilidade de cada modelo) não

responderem adequadamente, foi iniciada uma segunda fase do trabalho com a finalidade de definir modelos alternativos fiáveis de avaliação do risco de crédito dos clientes da empresa nacional.

O processo de modelação passou por diversas etapas, tendo-se utilizado uma bateria de 28 rácios económicos e financeiros e outros 3 indicadores financeiros e testado diferentes “zonas” de segmentação, por volume de negócios dos clientes.

Através da metodologia logit e da técnica forward stepwise (Wald), obtiveram-se dois modelos para a avaliação do risco de crédito de clientes, que por apresentarem um melhor ajustamento e precisão de entre os vários testes, foram os selecionados.

De seguida, apresenta-se uma síntese relativa aos dois modelos inicialmente testados e os dois modelos estimados.

■ Modelos Testados

Modelo1	$0,12 \cdot \text{ROA} + 0,03 \cdot \text{RAI} + 0,06 \cdot \text{ROE} + 0,06 \cdot \text{Marg. Bruta Vendas} + 0,03 \cdot \text{Rend. Líquida Vendas} + 0,12 \cdot \text{Remun. Cap. Alheios} + 0,13 \cdot \text{Garantia} + 0,07 \cdot \text{Aut. Financeira} + 0,06 \cdot \text{Peso End. LP} + 0,07 \cdot \text{Est. Financeira} + 0,04 \cdot \text{Liquidez Geral} + 0,06 \cdot \text{Liquidez Imediata} + 0,15 \cdot \text{Vendas}$
Modelo2	$0,13 \cdot \text{ROA} + 0,03 \cdot \text{RAI} + 0,06 \cdot \text{ROE} + 0,06 \cdot \text{Marg. Bruta Vendas} + 0,04 \cdot \text{Rend. Líquida Vendas} + 0,13 \cdot \text{Remun. Cap. Alheios} + 0,08 \cdot \text{Garantia} + 0,05 \cdot \text{Aut. Financeira} + 0,03 \cdot \text{Peso End. LP} + 0,04 \cdot \text{Est. Financeira} + 0,04 \cdot \text{Liquidez Geral} + 0,06 \cdot \text{Liquidez Imediata} + 0,25 \cdot \text{Vendas}$

■ Modelos Estimados

ModeloI	$-0,315 + 5,859 \cdot \text{Remun. Cap. Alheios} - 0,12 \cdot \text{Solvabilidade} + 1,817 \cdot \text{Peso End. LP} - 0,094 \cdot \text{Vendas} - 0,333 \cdot \text{Rácio Fundo Manéio} + 1,534 \cdot \text{Est. Endividamento}$
ModeloII	$-0,952 + 0,516 \cdot \text{Peso Endiv. LP} - 0,126 \cdot \text{Liquidez Geral} + 1,346 \cdot \text{Endividamento} + 0,072 \cdot \text{Cob. Ativo Fixo Tangível}$

Ambos os modelos estimados utilizam um menor número de variáveis do que os iniciais, esta menor necessidade de informação pode ser apontada como um aspecto positivo.

Enquanto os modelos testados utilizam uma maior diversidade de tipologias de rácios (rendibilidade, liquidez, alavancagem, funcionamento), os modelos estimados são constituídos essencialmente por dois tipos de rácios, os de alavancagem e de liquidez.

Este aspecto pode considerar-se limitativo da avaliação económica e financeira das empresas, uma vez que não contempla todas as suas vertentes.

		Modelos Testados	
Mod. 1: Clientes VN 8-100 Milhões€		Mod. 2: Clientes VN >100 Milhões€	
Observado	Previsto		Taxa Acerto
	Com Risco	Sem Risco	
Com Risco	100%	0%	100%
Sem Risco	100%	0%	0%
		45,9%	
Mod. I: Clientes VN < 1 Milhão€		Mod. II: Clientes VN > 1 Milhão€	
Observado	Previsto		Taxa Acerto
	Com Risco	Sem Risco	
Com Risco	78%	23%	78%
Sem Risco	57%	43%	43%
		61,1%	
Observado	Previsto		Taxa Acerto
	Com Risco	Sem Risco	
Com Risco	61%	39%	61%
Sem Risco	36%	64%	64%
		62,3%	
Modelos Estimados			

Através da leitura das matrizes de classificação, facilmente se conclui que os dois modelos estimados possuem maior taxa de acerto do que os testados. Embora o Modelo2 até apresente uma taxa geral de acerto superior à dos modelos estimados, e a taxa de acerto dos casos de incumprimentos seja de 100% nos dois modelos testados, verifica-se que a taxa de acerto dos casos ditos normais de cumprimento é de 0%, o que significa que ambos os modelos não conseguem prever nenhum destes casos, o que leva a crer que os critérios são desajustados ou não se encontram harmonizados de modo a distinguir os sinais de cumprimento e de incumprimento das empresas que fazem parte da amostra. Considera-se que os modelos testados, carecem de ajustamentos ou de uma nova calibração, com base na amostra de clientes atual.

Os dois modelos estimados e selecionados não apresentam taxas de acerto muito elevadas, o que se atribui em grande parte, aos critérios utilizados na classificação dos clientes da amostra como clientes “com risco” e clientes “sem risco”, e cujos constrangimentos convém reforçar.

Alguns clientes classificados como “Com Risco”, sobretudo os que possuem dívida vencida inferior a 6 meses, podem efetivamente não possuir problemas financeiros que levem a uma ruptura definitiva nos pagamentos, na realidade, pode ter ocorrido que os mesmos ainda não tenham tido oportunidade de efetuar algum pedido ou reclamação, ou mesmo, que estes ainda não estejam atualizados na base de dados, por outro lado,

existem aqueles clientes, normalmente grandes clientes, que sistematicamente pagam após a data de vencimento.

Por outro lado, admite-se a possibilidade de alguns clientes classificados como “Sem Risco”, apesar de manterem os seus pagamentos em dia, apresentarem já alguns problemas financeiros e de liquidez espelhados na sua contabilidade.

O facto da energia, nomeadamente o gás e a eletricidade constituírem factores de produção cruciais e indispensáveis em qualquer atividade económica, leva a crer que tendencialmente, uma empresa com dificuldades financeiras e de liquidez, deixará em primeiro lugar de pagar a outros fornecedores e tentará manter os pagamentos ao fornecedor de energia em dia. Esta perceção deixa antever situações de incumprimento graves, no sentido de que, o cliente deixará de pagar as faturas de energia, numa fase em que a sua situação financeira já se encontra bastante debilitada, resultando em casos de insolvência ou mesmo de falência, tornando-se muito difícil recuperar os créditos.

Ocorre, por vezes, este tipo de clientes, entrar em contacto com o fornecedor de energia, no sentido de efetivar um plano de pagamentos que consiga suportar. Este tipo de contactos podem constituir, em certa medida, um indicador importante na probabilidade do cliente em causa vir a entrar em incumprimento.

Caso estes constrangimentos não se verificassem, decerto a taxa de acerto nos dois modelos estimados seria superior.

Relativamente a estes constrangimentos, propõe-se em alternativa, a utilização de uma classificação da qualidade do crédito dos clientes, fornecida por entidade externa à empresa. No entanto, e por representar um custo adicional, esta metodologia não foi aplicada.

Por outro lado, o facto da carteira de clientes nacional ser constituída maioritariamente por pequenas empresas, levou a que se baixasse o valor do VN, para segmentar os clientes, que após vários testes, se decidiu situar em um milhão de euros.

Se por um lado não faz sentido baixar mais o limite de VN que define a segmentação, porque não se vê qualquer vantagem numa maior segmentação da carteira, por outro lado,

é possível que com a ampliação do mercado, se justifique uma deslocação para cima da do limite de VN, ou uma nova segmentação mais ao nível das grandes empresas.

5.2. Proposta de Melhorias

Relembrem-se os constrangimentos inerentes à avaliação do risco de crédito de uma empresa, atendendo exclusivamente à sua informação financeira, em que a utilização de contabilidade criativa ou mesmo a não disponibilização da informação financeira pelas empresas, são apenas alguns exemplos.

Conforme refere Platt (1985) os rácios deverão ser compatíveis, de forma a tornar a comparação entre empresas possível. A impossibilidade de detectar a existência de sazonalidade da actividade ou a alteração nos métodos contabilísticos utilizados e a possível ausência de termo de comparação quer em termos de dados reais setoriais e da empresa, quer em termos de valores empíricos, são aspectos a considerar.

A avaliação do risco de crédito de clientes, por intermédio da análise da situação económica e financeira de uma empresa revela-se uma componente de importância comprovada, porém, outros aspectos não deverão ser descurados na avaliação da qualidade do crédito do cliente.

Os modelos de avaliação do risco de crédito de clientes devem ser capazes de prever, pelo menos com um ano de antecedência, o evento de incumprimento do cliente. Pois bem, se por um lado, a metodologia dos rácios espelha a situação económica e financeira do cliente empresarial, permitindo perceber o estado de solvência e a sua capacidade de cumprir as obrigações financeiras que assume, também é verdade que existem factores exógenos que impactam diretamente na situação financeira das empresas que não podem ser descurados e que nem sempre têm um reflexo automático nas contas das mesmas.

Embora não tenha sido objetivo deste trabalho, avaliar e determinar quais e qual a significância desses factores na avaliação do risco de crédito de clientes, considera-se que os modelos “quantitativos” estimados podem ser complementados com uma componente

de avaliação da situação macroeconómica nacional e internacional, sobretudo no caso dos clientes exportadores, e uma componente setorial, através de diagnósticos atualizados do risco de setor da actividade a que a empresa pertence, bem como daqueles com que ela mais se relaciona, e dos quais os seus resultados mais dependem.

Os modelos podem ainda ser melhorados com informação gerada no conhecimento e do gestor de cliente, porque este já contactou com o cliente ou já visitou as suas instalações, e muito provavelmente, até conhece os seus stakeholders (concorrentes, fornecedores, clientes...).

A inexistência de uma base de dados de morosos também pode ser apontada como um constrangimento, no entanto, esta lacuna poderá ser colmatada, à semelhança do que ocorre em outros setores de negócio, como o bancário e o das telecomunicações, pela criação de uma base dados de morosos do setor energético, que seja partilhada por todas as comercializadoras de energia.

Convém reforçar, a relevância da empresa possuir uma segmentação de clientes por risco de crédito. Este tipo de segmentação potencia uma maior facilidade e eficácia, não só no delineamento da política de pricing, como também, na orientação da estratégia de marketing e das campanhas de angariação de clientes. Já para não falar, no aspecto financeiro e contabilístico, uma vez que o processo de provisionamento de clientes poderá tornar-se mais fácil, se os clientes forem provisionados de acordo com a sua classificação de risco de crédito.

O estudo sobre a definição do critério ou dos critérios de segmentação da carteira de clientes por Risco de Crédito, necessita ser aprofundado, até mesmo com o recurso a uma análise de benchmarking, porém, os resultados obtidos no trabalho elaborado paralelamente a este, sobre a exposição atual da empresa ao risco de incumprimento de clientes, deixam antever algumas conclusões interessantes, nomeadamente ao nível de uma eventual segmentação do risco por atividade económica do cliente.

Acredita-se que os modelos estimados poderão sofrer uma melhoria na sua capacidade de previsão, através da introdução de outras variáveis de cariz qualitativo, nomeadamente as sugeridas anteriormente, porém, considera-se que a componente quantitativa dos modelos estimada, assim como os requisitos levantados, são um contributo importante para o desenvolvimento de futuros trabalhos.

BIBLIOGRAFIA

- Altman, E.I. (1968), "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy ", *Journal of Finance*, 23, 4, 589-609.
- Altman, E.I. & Haldeman, R.G. & Narayanan, P. (1977), "Zeta Analysis. A new model to identify bankruptcy of corporations", *Journal of Banking and Finance*, 1, 29-54.
- Altman, E.I. & Marco, G. & Varetto, F. (1994), "Corporate Distress Diagnosis: Comparing using Linear Discriminant Analysis and Neural Networks (the italian experience)", *Journal of Banking and Finance*, 18, 505-529.
- Altman, E. I. & Sabato, Gabriele (2007), "Modeling Credit Risk for SMEs": Evidence from the US Market; *Journal of Financial Services Research*, 3, 332-357.
- Argenti, J. (1976), "Corporate Collapse: The Causes and Symptoms", London, McGraw-Hill.
- Altuzarra, C. (2011). "Crise obriga exportadoras a diversificar mercados". *Diário Económico*, 5163, pp.16.
- Barniv, R. & J. McDonald (1999) "Review of Categorical Models for Classification; Issues in Accounting and Finance", *Review of Quantitative Finance and Accounting* 13, 39-62.
- Bardos, M. & Zhu, W.H. (1997), "Comparaison de l'Analyse Discriminante Linéaire et des Réseaux de Neurones, Application à la Détection de Défaillances d'Entreprises », *Revue de Statistique Appliquée*, 45, 65-92.
- Beaver, W. (1966), "Financial Ratios as Predictors of Failure", *Journal of Accounting Research* (Supplement), 4, pp 71-102.
- Bessis, J. (1998), "Risk Management in Banking", New York, John Wiley & Son Ltd.
- Black, F. & Scholes, M. (1973), "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, 81, 637-659.
- Boudoukh, J. & A. Saunders (1998), "Advanced Topics in Market Risk Measurement and Management", Seminário organizado pela *Stern School of Business*, New York Business, Fevereiro 1998.
- Bonfim, D. (2007), "Factores determinantes do risco de crédito: O contributo de características das empresas e da envolvente macroeconómica", Banco de Portugal. Relatório de Estabilidade Financeira 2006.
- Borooah, V. (2002), "Logit and Probit: Ordered and Multinomial Models", Quantitative Applications in the Social Sciences, Series N.º 138, California, Sage Publications: Thousand Oaks.
- Brealey, Richard A. & Myers, Stewart C. (1998), *Princípios de Finanças Empresariais*, McGraw-Hill de Portugal.
- Carvalho das Neves J. (2005), "Análise Financeira: Técnicas Fundamentais" (*Edição Actualizada*), Lisboa, Texto Editores.
- Carvalho das Neves, J. & Andrade e Silva, J. (1998), "Modelos de Análise do Risco de Incumprimento à Segurança Social", Centro de Estudos e Documentação Europeia do Instituto Superior de Economia e

Gestão da Universidade Técnica de Lisboa e Fundação para a Ciência e Tecnologia – Projecto SSPS/S/AEA/2617/96.

Casey, C.J. & Baztczak, N.J. (1985), “Cash-Flow It’s Not The Bottom Line”, *Harvard Business Review*, 4, 60-66.

Charnet, R. et al (1999), “Análise de modelos de regressão linear com aplicações”, Campinas:Editora da Unicamp.

Chorão, L.A.R. (2005), “Logit vs Redes Neuronais Artificiais: Um exemplo”, Tese de Mestrado em Estatística e Gestão de Informação, ISEGI-UNL.

Coface (2011), Estudo de Insolvências de Empresa – Portugal – 1º Trimestre 2010/2011, acedido em Junho 2011 em www.coface.pt

Cook, R. A. & Nelson, J. L. (1988), “A Conspectus of Business Failure Forecasting”, consultado em Junho de 2011 em <http://www.sbaer.uca.edu/research/sbida/1988/PDF/22.pdf>

Crook, Jonathan & Edelman, David & Thomas, Lyn (2007), “Recent Developments in Consumer Credit Risk Assessment”, *European Journal of Operational Research*, 183, pp. 1447-1465.

Deakin, E. B. (1972), “A Discriminant Analysis of Predictors of Business Failure”, *Journal of Accounting Research*, 1, 167-179.

Falkenstein, Eric (2000), “Risk Calc™ For Private Companies: Moody’s Default Model”, Moddy’s Investors Service.

Ferrando, M. & Blanco, F. (1998). “La Prevision del Fracaso Empresarial en la Comunidad valenciana: Aplicacion de los Modelos Discriminante y Logit”, *Revista Española de Financiación y Contabilidad*.

Fons, J.S. (1994), “Using Default Rates to Model the Term Structure of Credit Risk”, *Financial Analyst Journal*, Set-Oct., 25-32.

Gabas, F. (1990), “Técnicas Actuales de Análisis Contable - Evaluacion de la Solvencia Empresarial”, Instituto de Contabilidad y Auditoría de Cuentas, Madrid.

Gabinete de Planeamento, Estratégia, Avaliação e Relações Internacionais (2011). Previsões de primavera de 2011 da CE, FMI e OCDE. Consultado em Junho de 2011 em: <http://www.gpeari.min-financas.pt/analise-economica/publicacoes/ficheiros-do-bmep/2011/maio/destaques/Previsoes-de-primavera-de-2011-da-CE-FMI-e-OCDE.pdf/view>

Gamelas, Emanuel Freire Torres (2005), “Falências de Empresas em Portugal”, Tese de Doutoramento em Economia Financeira e Contabilidade, Instituto Superior de Ciências do Trabalho e da Empresa.

Hair, J. & Anderson, R. & Tatham, R. & Black, W. (2005), “Análise Multivariada de Dados”. 5ª. ed. Porto Alegre: Bookman.

Hosmer, D., Lemeshow, S. (1998), “Applied Logistic Regression”, New York: John Wiley & Sons, Inc.

Johnston, J. & Dinardo, J. (1997), *Econometric Methods*, New York, McGraw-Hill.

Klimasauskas, C. C. (1993), “Applying Neural Networks”, in *Neural Networks in Finance and Investing*, Chicago, Trippi and Turban eds, Probus Publishing Company.

Laitinen, E. (1999), “Predicting a Corporate Credit Analyst’s Risk Estimate by Logistic and Linear Models”, *International Review of Financial Analysis* 8, 97-12.

Lewis, Edward M. (1992), "An Introduction to Credit Scoring", Vol. Seconde Edition. San Rafael, California: Fair, Isaac and Co.,Inc.

Lo, A. W. (1986), "Logit Versus Discriminant Analysis: A Specification Test and Application to Corporate Bankruptcy", *Journal of Econometrics*, 31, 151-178.

Loureiro, João. (2011). "Cada vez mais famílias deixam de pagar bens comprados a crédito". *Diário Económico*, 5009, pp. 10-11.

Manso, José Pires (2010), "Insolvências de Empresas", Observatório para o Desenvolvimento Económico e Social.

Meadows et al. (1972), "Limits to Growth", New York, Potomac Associates.

Merton, R.C. (1974), "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates", *Journal of Finance*, 29, 449-470.

Mora, A. (1996). "El Contenido Informativo de los Datos Contables para las Decisiones de Inversion". Instituto de Contabilidad y Auditoría de Cuentas Madrid.

Nogueira Leite, A. (2009). "Falências em Portugal aumentam 80%". *Diário Económico*, 4608, pp.6.

Ohlson, J. (1980), "Financial ratios and the probabilistic prediction of bankruptcy", *Journal of Accounting Research*, 18, 109-131.

Pampel, F. (2000), "Logistic Regression: A Primer", Quantitative Applications in the Social Sciences, Series N.º 132. Thousand Oaks: Sage Publications.

Platt, H. & M. Platt (1990), "Development of a class of stable predictive variables: the case of bankruptcy prediction", *Journal of Business Finance and Accounting*, 17, 31-51.

Platt, H. D. (1985), "Why Companies Fail: Strategies for Detecting Avoiding and Profiting from Bankruptcy", Massachusetts: Lexington Books.

Raymond, Anderson (2007), "The Credit Scoring Toolkit Theory and Practice for Retail Credit Risk Management and Decision Automation", New York: OXFORD University Press Inc.

Richardson, F. & Kane, G. & Lobingier, P. (1998), "The Impact of Recession on the Prediction of Corporate Failure", *Journal of Business, Finance & Accounting*, 25.

Saunders, A. (1997), "Financial Institutions Management: A Modern Perspective", 2ª Ed., New York: Irwin McGraw-Hill.

Sharma, S. (1996), "Applied multivariate techniques", New York: John Wiley & Sons.

Sheppard, J. P. (1994), "The Dilemma of Matched Pairs and Diversified Firms in Bankruptcy Prediction Models", *The Mid-Atlantic Journal of Business*, 30 (March), 9-25.

Silva, José Pereira da, 2004, "Gestão e Análise de Risco de Crédito", São Paulo :Editora Atlas.

SPSS (2003), "*SPSS Base 12.0 User's Guide*", Chicago: IL: SPSS, Inc.

Srinivasan, V. & Kim Y.H (1987), "Credit Granting: A comparative Analysis of Classifications Procedures", *The Journal of Finance*, 42, 665-683.

Stickney, C., P. (1996), "Financial Reporting and Statement Analysis", 3rd Edition, Ft. Worth, TX: The Dryden Press.

Thomas, Lyn C. (2009), "Consumer Credit Models: Pricing, Profit and Portfolios", New York: Oxford University Press Inc.

Thomas, Lyn C. & David B. Edelman & Jonathon N. Crook, (2002), "Credit Scoring and Its Applications", Siam Monographs on Mathematical Modeling and Computation .

Tomas, J. & Amat, O. & Esteve, M. (1999), "Como Analizan las Entidades Financieras a sus Clientes", Editorial Gestión 2000, Barcelona.

Trippi, R.R. & Turban E. (1993), "Neural Networks in Finance and Investing", Probus Publishing Company, Chicago.

Zavgren, C.(1985), "Assessing the Vulnerability to Failure of American Industrial Firms: A Logistic Analysis", *Journal of Business Finance and Accounting*, 12, 19-45

Zavgren C. (1983), "The Prediction of Corporate Failure: The State of the Art", *Journal of Accounting Literature*, vol. 2, 1-38.

Anexo 1 – Análise Discriminante

Um modelo discriminante é uma técnica que tenta encontrar um sistema para classificar os sujeitos numa série de grupos bastante homogêneos a nível interno (entre cada um dos sujeitos) e bastante heterogêneos a nível externo (entre cada um dos grupos).

O algoritmo mais adequado é a análise de clustering, sempre que não exista de modo pré-determinado uma variável de agrupamento (isto é, uma variável de classificação). O objetivo do algoritmo de clustering é criar uma variável de classificação a partir do cálculo de distâncias (isto é, semelhanças ou diferenças) ótimas entre os sujeitos.

Uma vez obtida a variável classificadora ou de segmentação, chega o momento de utilizar essa variável no sentido de tentar reclassificar os grupos, o que se conhece pelo nome técnico de discriminar.

Daí vem o nome do outro algoritmo de classificação – Análise Discriminante. A Análise Discriminante não pretende criar grupos de semelhança, utilizando os já existentes (criados a partir de um algoritmo de *clustering* – análise de conglomerados) para oferecer uma análise da capacidade das variáveis da Análise Discriminante (chamadas, neste algoritmo, “variáveis discriminantes”) para classificar ou diferenciar (discriminar) entre os n sujeitos dos k grupos. Deste modo, obter-se-ia um esquema de classificação.

O modelo de Análise Discriminante vem definido por uma combinação linear chamada Função Discriminante:

$$D = B_0 + B_1 X_1 + B_2 X_2 + \dots + B_k X_k + E_{ij}$$

Onde:

- X_1, X_2, \dots, X_k - variáveis discriminantes,
- B_0 - constante,
- B_1, B_2, \dots, B_k - coeficientes discriminantes não estandardizados de variáveis,
- E_{ij} - variável aleatória que representa a parte do erro do modelo,
- k – número de variáveis discriminantes.

O valor da Função Discriminante para um indivíduo calcula-se substituindo na equação discriminante os valores das Variáveis Discriminantes, como em qualquer modelo de regressão. Ao valor assim obtido chama-se Pontuação Discriminante (Score).

Anexo 2 – Modelo Z-Score de Altman

• Rácios Financeiros Utilizados:

$X_1 = \frac{\text{Fundo de Maneto}}{\text{Activo Total}}$	$X_4 = \frac{\text{Valor de Mercado dos Capitais Próprios}}{\text{Passivo Total}}$
$X_2 = \frac{\text{Resultados Retidos}}{\text{Activo Total}}$	$X_5 = \frac{\text{Vendas}}{\text{Activo Total}}$
$X_3 = \frac{\text{Resultados antes de juros e impostos}}{\text{Activo Total}}$	

A função discriminante obtida por Altman foi a seguinte:

$$\hat{Z} = 1,2X_1 + 1,4X_2 + 3,3X_3 + 0,6X_4 + 0,999X_5$$

Aplicando o modelo construído à amostra, verifica-se um erro tipo I de 6% e 28%, um e dois anos antes da falência, respetivamente, o que significa que o modelo classificou empresas que na realidade se encontram falidas, como estando na situação de não falidas.

Por outro lado, verifica-se um Erro Tipo II de 3% e 6%, um e dois anos antes da falência, respetivamente, o que significa que o modelo classificou empresas que na realidade não se encontram falidas, como estando na situação de falidas.

Quadro 32 - Matriz de classificação (um ano antes da falência)

Realidade	Previsão	
	Falida	Não Falida
Falida	-	6%
Não Falida	3%	-

Fonte: Altman, 1993

Quadro 33 - Matriz de classificação (dois anos antes da falência)

Realidade	Previsão	
	Falida	Não Falida
Falida	-	28%
Não Falida	6%	-

Fonte: Altman, 1993

O modelo foi novamente estimado, utilizando o valor dos capitais próprios em vez do valor de mercado.

A nova função discriminante estimada por Altman foi:

$$\hat{Z} = 0,717X_1 + 0,847X_2 + 3,107X_3 + 0,42X_4 + 0,998X_5$$

Anexo 3 – Análise Descritiva das Variáveis

Quadro 34 – Análise Descritiva – Clientes “Sem Risco”

Descriptive Statistics									
	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
ROA	981	-2,43	3,50	0,01	0,18	3,75	0,08	177,92	0,16
RAI	981	0,00	19,00	7,61	5,42	-0,47	0,08	-1,32	0,16
ROE	981	-79,27	41,41	0,00	3,14	-14,52	0,08	451,07	0,16
Margem Bruta Vendas	981	-12,00	10,00	-0,04	0,66	-4,97	0,08	200,89	0,16
Rend Líquida Vendas	981	-12,00	10,00	-0,05	0,66	-5,10	0,08	205,67	0,16
Remun Capitais Alheios	981	0,00	0,42	0,03	0,03	3,77	0,08	27,23	0,16
Solvabilidade	981	0,10	421,61	3,54	16,53	19,45	0,08	444,77	0,16
Auton Financeira	981	-0,90	420,61	2,54	16,53	19,45	0,08	444,77	0,16
Peso End LP	981	0,00	1,00	0,24	0,27	0,92	0,08	-0,29	0,16
Estrutura Financeira	981	-230,85	955,50	3,68	37,07	18,23	0,08	452,77	0,16
Liquidez Geral	981	0,00	312,45	3,37	12,63	18,31	0,08	405,67	0,16
Liquidez Imediata	981	-0,65	63,36	0,82	3,20	11,33	0,08	177,37	0,16
Vendas	981	0,00	21,00	14,38	2,31	-2,45	0,08	15,28	0,16
Rácio do FM	981	-5,08	1,00	0,17	0,45	-4,42	0,08	40,46	0,16
Rotação Ativo	981	0,00	18,00	1,12	1,08	5,62	0,08	70,84	0,16
Rotação do FM	981	-4860,00	1541,00	-1,37	170,67	-22,67	0,08	680,61	0,16
Endividamento	981	0,00	9,92	0,64	0,55	8,21	0,08	106,67	0,16
Estrutura do Cap Permanente	981	-101,23	156,48	0,73	9,08	4,70	0,08	143,93	0,16
Cobertura Ativo Fixo Tangível	981	-51,35	267,94	0,86	9,29	24,98	0,08	707,40	0,16
Vendas Por Meios Financeiros Líquidos	978	-2,00	42016,00	201,35	1936,40	20,03	0,08	421,93	0,16
Vendas Por Contas a Receber	981	0,00	237479,00	303,60	7742,97	29,65	0,08	902,47	0,16
Rotação das Existências	981	0,00	491135,00	748,61	16961,55	26,45	0,08	735,12	0,16
Rotação do Capital Próprio	981	-139,00	1369,00	5,00	47,38	25,08	0,08	707,97	0,16
Cobertura dos Custos Financeiros	981	-15623,31	591,13	-24,84	538,43	-25,92	0,08	729,58	0,16
Cobertura do Ativo por Resultados Transitados	981	-10,14	0,85	-0,10	0,74	-8,74	0,08	92,66	0,16

Dívidas a Terceiros C/P sobre capital próprio	981	-116,79	955,50	2,78	33,64	23,91	0,08	662,42	0,16
Fundo de Maneio sobre capital próprio	981	-415,18	67,78	-0,19	15,31	-21,56	0,08	562,84	0,16
Estrutura do Endividamento	981	0,00	1,00	0,64	0,27	-0,50	0,08	-0,79	0,16
Rendibilidade Acumulada	981	-10,12	1,07	0,10	0,75	-8,14	0,08	85,88	0,16
Ativos Tangíveis	981	0,00	22,97	14,56	2,00	-1,91	0,08	16,68	0,16
Meios Financeiros Líquidos sobre o Ativo	981	-0,47	0,93	0,11	0,16	2,09	0,08	4,95	0,16

Quadro 35 - Análise Descritiva – Clientes “Com Risco”

Descriptive Statistics									
	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
ROA	981	-2,47	0,58	-0,02	0,16	-5,70	0,08	63,25	0,16
RAI	981	0,00	20,00	6,65	5,57	-0,12	0,08	-1,48	0,16
ROE	981	-57,56	14,83	-0,20	3,07	-13,07	0,08	231,77	0,16
Margem Bruta Vendas	981	-6733,00	102,00	-7,71	216,18	-30,82	0,08	958,60	0,16
Rend Líquida Vendas	981	-6733,00	101,00	-7,75	216,24	-30,79	0,08	957,52	0,16
Remun Capitais Alheios	981	0,00	1,92	0,04	0,08	18,40	0,08	390,12	0,16
Solvabilidade	981	0,00	143,30	2,02	6,03	19,67	0,08	422,90	0,16
Auton Financeira	981	-0,87	142,30	1,02	6,03	19,68	0,08	422,96	0,16
Estrutura Financeira	981	-927,69	1809,04	6,49	77,17	12,04	0,08	333,94	0,16
Liquidez Geral	981	0,00	577,27	2,37	18,63	30,10	0,08	928,50	0,16
Liquidez Imediata	981	-0,07	33,53	0,40	1,63	12,26	0,08	206,68	0,16
Vendas	981	0,00	22,00	13,91	3,23	-2,49	0,08	9,40	0,16
Rácio do FM	981	-5,88	1,00	0,02	0,47	-4,27	0,08	41,20	0,16
Rotação do Ativo	981	0,00	17,00	1,11	1,35	5,78	0,08	55,08	0,16

Rotação do FM	981	-7565,00	1025,00	-6,18	261,78	-25,57	0,08	721,33	0,16
Endividamento	981	0,00	7,51	0,78	0,52	6,37	0,08	63,25	0,16
Estrutura do Cap Permanente	981	-179,68	1200,31	2,82	40,47	26,63	0,08	785,80	0,16
Cobertura Ativo Fixo Tangível	981	-454,37	393,42	1,23	21,99	-0,27	0,08	332,90	0,16
Vendas Por Meios Financeiros Líquidos	978	-122,00	15359,00	222,36	1013,49	10,79	0,08	135,31	0,16
Vendas Por Contas a Receber	981	0,00	280167,00	300,02	8944,92	31,32	0,08	980,85	0,16
Rotação das Existências	981	0,00	55924,00	159,00	1985,19	24,19	0,08	650,76	0,16
Rotação do Capital Próprio	981	-1372,00	1436,00	8,67	82,99	6,02	0,08	256,27	0,16
Cobertura dos Custos Financeiros	981	-6617,74	1610161,00	1632,98	51409,23	31,32	0,08	980,96	0,16
Cobertura do Ativo por Resultados Transitados	981	-101,70	0,79	-0,27	3,34	-28,82	0,08	874,34	0,16
Dívidas a Terceiros C/P sobre capital próprio	981	-862,98	559,32	3,32	46,97	-1,73	0,08	175,56	0,16
Fundo de Maneio sobre capital próprio	981	-446,68	393,71	0,10	24,61	-2,77	0,08	213,34	0,16
Estrutura do Endividamento	981	0,00	1,00	0,63	0,29	-0,47	0,08	-0,88	0,16
Rendibilidade Acumulada	981	-96,18	0,82	-0,13	3,17	-28,66	0,08	867,48	0,16
Ativos Tangíveis	981	0,00	22,40	14,44	2,25	-1,88	0,08	12,97	0,16
Meios Financeiros Líquidos sobre o Ativo	981	-0,05	0,91	0,08	0,14	3,05	0,08	10,35	0,16

Anexo 4 – Matriz de Correlações de Pearson

Quadro 36 – Matriz de Correlações de Pearson

	ROA	RAI	ROE	Margem Bruta Vendas	Rend Líquida Vendas	Remun Capitais Alheios	Solvabilidade	Auton Financeira	Estrutura Financeira	Liquidez Geral	Liquidez Imediata	Ln Vendas	Rácio do FM
ROA	1	,474**	,054*	,074**	,074**	-,055*	0,027	0,027	-0,004	-,117**	,056*	,099**	,220**
RAI		1	,098**	0,035	0,035	-0,008	0,025	0,025	-,049*	0,026	,046*	,301**	,220**
ROE			1	-0,004	-0,004	-0,003	0,005	0,005	-,231**	0,006	0,009	0,026	0
Margem Bruta Vendas				1	1,000**	0,013	0,005	0,005	0,003	0,005	0,006	,079**	,118**
Rend Líquida Vendas					1	0,013	0,005	0,005	0,003	0,004	0,006	,079**	,118**
Remun Capitais Alheios						1	-0,005	-0,005	-0,01	0,01	0,003	,060**	0,044
Solvabilidade							1	1,000**	-0,011	,275**	,562**	-,125**	,065**
Auton Financeira								1	-0,011	,275**	,562**	-,125**	,065**
Estrutura Financeira									1	-0,009	-0,015	-0,028	-0,022
Liquidez Geral										1	,370**	-0,039	,144**
Liquidez Imediata											1	-,065**	,197**
Vendas												1	,067**
Rácio do FM													1
Rotação do Ativo													0,003
Rotação do FM													0,014
Endividamento													0,526
Estrutura do Cap Permanente													0
Cobertura do Ativo Fixo Tangível													0,987
Vendas Por Meios Financeiros Líquidos													0,444
Vendas Por Contas a Receber													-0,003
Rotação das Existências													0,88
Rotação do Capital Próprio													-0,024
Cobertura dos Custos Financeiros													0,288
Cobertura do Ativo por Resultados Transistados													0,021
Dívidas a Terceiros C/P sobre capital próprio													0
Fundo de Maneio sobre capital próprio													0,151
Estrutura do Endividamento													0,002
Rendibilidade Acumulada													0
Ativos Tangíveis													0,006
Meios Financeiros Líquidos sobre o Ativo													,271**

A tabela apresenta o grau de associação linear entre as variáveis medido através do coeficiente de Correlação R de Pearson.
 **, Correlation is significant at the 0.01 level (2-tailed); *, Correlation is significant at the 0.05 level (2-tailed).

	Rotação do Ativo	Rotação do FM	Endividamento	Estrutura do Cap. Permanente	Cobertura Ativo Fixo Tangível	Vendas Por Meios Financeiros Líquidos	Vendas Por Contas a Receber	Rotação das Existências	Rotação do Capital Próprio	Cobertura dos Custos Financeiros	Cobertura do Ativo por Resultados Transitados
ROA	-0,011	0,012	-,390**	-0,004	-0,002	0,006	0	0,013	0	0,009	,094**
	0,611	0,605	0	0,858	0,941	0,792	0,97	0,567	0,992	0,69	0
RAI	,107**	0,023	-,236**	-0,041	-0,037	0,023	0,002	-0,012	-0,041	0,014	,090**
	0	0,316	0	0,071	0,106	0,309	0,913	0,588	0,07	0,545	0
ROE	0,01	0,004	-0,01	-,217**	-0,024	0,002	0,007	0,002	-,167**	0,001	0,031
	0,664	0,864	0,654	0	0,279	0,914	0,758	0,923	0	0,954	0,174
Margem Bruta	0,023	0	-,073**	0,002	0,001	0,003	0,001	0,001	0,003	0,001	0,019
Vendas	0,307	0,985	0,001	0,941	0,961	0,88	0,968	0,965	0,91	0,977	0,403
Rend Líquida	0,023	0	-,073**	0,002	0,001	0,003	0,001	0,001	0,003	0,001	0,019
Vendas	0,305	0,985	0,001	0,941	0,961	0,879	0,967	0,967	0,909	0,977	0,403
Remun Capitais	,061**	-0,001	-,060**	-0,005	-0,006	-0,005	-0,013	-0,007	-0,003	-0,012	0,027
Alheios	0,007	0,948	0,008	0,809	0,775	0,827	0,575	0,758	0,877	0,581	0,235
Solvabilidade	-,048*	0,003	-,146**	-0,008	-0,005	-0,012	0	-0,004	-0,012	0,019	0,012
	0,035	0,878	0	0,726	0,811	0,607	0,977	0,86	0,607	0,408	0,594
Auton Financeira	-,048*	0,003	-,146**	-0,008	-0,005	-0,012	0	-0,004	-0,012	0,019	0,012
	0,035	0,878	0	0,726	0,811	0,606	0,977	0,86	0,607	0,408	0,594
Estrutura Financeira	0,001	0	0,023	,759**	,552**	0,001	0	0	,847**	-0,002	-0,034
	0,976	0,99	0,316	0	0	0,962	0,974	0,975	0	0,934	0,132
Liquidez Geral	-0,005	0,003	,086**	-0,005	-0,007	-0,012	0	-0,005	-0,01	0,007	-0,021
	0,832	0,881	0	0,818	0,759	0,597	0,975	0,831	0,65	0,76	0,355
Liquidez Imediata	0,024	0,005	-,170**	-0,011	-0,009	-0,032	0,002	-0,004	-0,014	,063**	0,024
	0,282	0,825	0	0,614	0,695	0,156	0,928	0,845	0,541	0,005	0,282
Vendas	,209**	0,017	-,055*	0,007	-,085**	,059**	-0,023	0,003	0,015	-,113**	,155**
	0	0,462	0,015	0,77	0	0,009	0,315	0,898	0,516	0	0
Rácio do FM	0,001	0,014	-,676**	0	-,046*	-0,017	-0,003	-0,028	-0,024	0,021	,185**
	0,957	0,526	0	0,987	0,041	0,444	0,88	0,208	0,288	0,363	0
Rotação do Ativo	1	0,023	,083**	-0,016	-0,004	-0,004	0,03	-0,009	,079**	-0,02	-0,002
		0,316	0	0,491	0,858	0,842	0,185	0,687	0	0,365	0,931
Rotação do FM	0,023	1	-0,011	0	0	0,005	0,001	0,002	0,002	0	0
	0,316	0,634	0,634	0,976	0,979	0,821	0,965	0,919	0,915	0,987	0,971
Endividamento	,083**	-0,011	1	0,019	0,015	0,007	-0,005	0,003	0,023	-0,028	-,232**
	0	0,634	0,401	0,506	0,765	0,815	0,895	0,308	0,222	0	0
Estrutura do Cap Permanente	-0,016	0	0,019	1	-0,001	0,002	-0,003	-0,002	,490**	-0,001	0
	0,491	0,976	0,401	0,955	0,913	0,908	0,936	0	0,949	0,969	0,969
Cobertura Ativo Fixo Tangível	-0,004	0	0,015	-0,001	1	-0,002	0,004	0,003	,660**	-0,001	-,046*
	0,858	0,979	0,506	0,955		0,946	0,843	0,911	0	0,957	0,042
Vendas Por Meios Financeiros Líquidos	-0,004	0,005	0,007	0,002	-0,002	1	-0,004	-0,004	0	-0,003	0,004
	0,842	0,821	0,765	0,913	0,946		0,842	0,854	0,985	0,893	0,85
Vendas Por Contas a Receber	0,03	0,001	-0,005	-0,003	0,004	-0,004	1	-0,001	0,003	0	0,003
	0,185	0,965	0,815	0,908	0,843	0,842		0,953	0,895	0,973	0,889
Rotação das Existências	-0,009	0,002	0,003	-0,002	0,003	-0,004	-0,001	1	-0,001	0	0
	0,687	0,919	0,895	0,936	0,911	0,854	0,953		0,956	0,97	0,979
Rotação do Capital Próprio	,079**	0,002	0,023	,490**	,660**	0	0,003	-0,001	1	-0,002	-0,033
	0	0,915	0,308	0	0	0,985	0,895	0,956		0,921	0,145
Cobertura dos Custos Financeiros	-0,02	0	-0,028	-0,001	-0,001	-0,003	0	0	-0,002	1	0,005
	0,365	0,987	0,222	0,949	0,957	0,893	0,973	0,97	0,921		0,81
Cobertura do Ativo por Resultados Transitados	-0,002	0	-,232**	0	-,046*	0,004	0,003	0	-0,033	0,005	1
	0,931	0,971	0	0,969	0,042	0,85	0,889	0,979	0,145	0,81	
Dívidas a Terceiros C/P sobre capital próprio	0,01	0,001	0,019	,378**	,777**	0	0	0	,846**	-0,002	-,050*
	0,659	0,972	0,394	0	0	0,985	0,983	0,993	0	0,944	0,026
Fundo de Maneio sobre capital próprio	-0,002	0,001	-0,01	,195**	-,831**	-0,003	-0,003	-0,002	-,472**	0	,047**
	0,926	0,948	0,665	0	0	0,904	0,906	0,919	0	0,985	0,036
Estrutura do Endividamento	,210**	0,002	-0,039	-,064**	0,027	-0,014	0,008	0,007	0,018	0,03	-0,018
	0	0,914	0,082	0,004	0,236	0,547	0,725	0,76	0,428	0,183	0,415
Rendibilidade Acumulada	-0,005	0,001	-,255**	-0,004	-,048*	0,003	0,004	-0,002	-0,038	0,007	,996**
	0,84	0,966	0	0,864	0,033	0,901	0,87	0,935	0,094	0,749	0
Ativos Tangíveis	-,213**	-0,002	-,104**	0,02	0	,054*	-,047*	0,007	-0,028	-0,043	,068**
	0	0,941	0	0,365	0,984	0,016	0,039	0,748	0,223	0,057	0,002
Meios Financeiros Líquidos sobre o Ativo	,215**	-0,002	-,187**	-0,027	-0,004	-,081**	0,018	-0,007	0,003	,074**	0,04
	0	0,942	0	0,239	0,863	0	0,418	0,767	0,892	0,001	0,08

A tabela apresenta o grau de associação linear entre as variáveis medido através do coeficiente de Correlação R de Pearson.

** . Correlation is significant at the 0.01 level (2-tailed); * . Correlation is significant at the 0.05 level (2-tailed).

	Dívidas a Terceiros C/P sobre capital próprio	Fundo de Maneio sobre capital próprio	Estrutura do Endividamento	Rendibilidade Acumulada	Ativos Tangíveis	Meios Financeiros Líquidos sobre o Ativo
ROA	-0,004 0,863	-0,002 0,913	0,038 0,094	,092** 0	,076** 0,001	,127** 0
RAI	-0,04 0,076	0,022 0,338	,078** 0,001	,090** 0	,125** 0	,134** 0
ROE	-,163** 0	-,097** 0	0,004 0,868	0,029 0,202	0,01 0,658	0,004 0,842
Margem Bruta Vendas	0,003 0,891	-0,002 0,92	-0,028 0,211	0,012 0,595	-0,009 0,706	0,012 0,599
Rend Líquida Vendas	0,003 0,891	-0,002 0,919	-0,028 0,21	0,012 0,596	-0,009 0,705	1962 0,012
Remun Capitais Alheios	-0,011 0,614	0,007 0,763	0,014 0,543	0,031 0,17	,050* 0,027	0,599 -0,003
Solvabilidade	-0,009 0,677	0,003 0,9	-0,016 0,474	0,022 0,333	-,156** 0	0,898 ,068**
Auton Financeira	-0,009 0,677	0,003 0,9	-0,016 0,472	0,022 0,333	-,156** 0	0,003 ,068**
Estrutura Financeira	,889** 0	-,396** 0	-0,011 0,637	-0,038 0,088	0 0,993	0,003 -0,008
Liquidez Geral	-0,009 0,695	0,009 0,688	-,132** 0	-0,007 0,749	-,109** 0	0,729 ,089**
Liquidez Imediata	-0,013 0,556	0,008 0,707	-,132** 0	0,042 0,064	-,149** 0	0 ,416**
Vendas	-,048** 0,035	,074** 0,001	0,034 0,136	,154** 0	,510** 0	0 -,066**
Rácio do FM	-0,032 0,151	,071** 0,002	-,230** 0	,206** 0	-,062** 0,006	0,004 ,271**
Rotação do Ativo	0,01 0,659	-0,002 0,926	,210** 0	-0,005 0,84	-,213** 0	0 ,215**
Rotação do FM	0,001 0,972	0,001 0,948	0,002 0,914	0,001 0,966	-0,002 0,941	0 -0,002
Endividamento	0,019 0,394	-0,01 0,665	-0,039 0,082	-,255** 0	-,104** 0	0,942 -,187**
Estrutura do Cap Permanente	,378** 0	,195** 0	-,064** 0,004	-0,004 0,864	0,02 0,365	0 -0,027
Cobertura Ativo Fixo Tangível	,777** 0	-,831** 0	0,027 0,236	-,048* 0,033	0 0,984	0,239 -0,004
Vendas Por Meios Financeiros Líquidos	0 0,985	-0,003 0,904	-0,014 0,547	0,003 0,901	,054* 0,016	0,863 -,081**
Vendas Por Contas a Receber	0 0,983	-0,003 0,906	0,008 0,725	0,004 0,87	-,047* 0,039	0 0,018
Rotação das Existências	0 0,993	-0,002 0,919	0,007 0,76	-0,002 0,935	0,007 0,748	0,418 -0,007
Rotação do Capital Próprio	,846** 0	-,472** 0	0,018 0,428	-0,038 0,094	-0,028 0,223	0,767 0,003
Cobertura dos Custos Financeiros	-0,002 0,944	0 0,985	0,03 0,183	0,007 0,749	-0,043 0,057	0,892 ,074**
Cobertura do Ativo por Resultados Transitados	-0,050* 0,026	,047* 0,036	-0,018 0,415	,996** 0	,068** 0,002	0,001 0,04
Dívidas a Terceiros C/P sobre capital próprio	1 0	-,703** 0	0,032 0,162	-,055* 0,016	-0,014 0,521	0,385 0,008
Fundo de Maneio sobre capital próprio	-,703** 0	1 0	-,047* 0,039	,050* 0,028	0,008 0,712	0,719 0,002
Estrutura do Endividamento	0,032 0,162	-,047* 0,039	1 0,018	-0,018 0,427	-,172** 0	0,921 ,073**
Rendibilidade Acumulada	-,055* 0,016	,050* 0,028	-0,018 0,427	1 0,001	,075** 0,001	0,001 ,056*
Ativos Tangíveis	-0,014 0,521	0,008 0,712	-,172** 0	,075** 0,001	1 0	0,014 -,212**
Meios Financeiros Líquidos sobre o Ativo	0,008 0,719	0,002 0,921	,073** 0,001	,056* 0,014	-,212** 0	0 1

A tabela apresenta o grau de associação linear entre as variáveis medido através do coeficiente de Correlação R de Pearson.

**. Correlation is significant at the 0.01 level (2-tailed); *. Correlation is significant at the 0.05 level (2-tailed).

Anexo 5 – Modelos Rejeitados

Quadro 37 – Matriz de Classificação – Modelo Sem Segmentação

		Predicted		
		Y		Percentage Correct
		0	1	
Observed				
Step 5 Y	0	592	386	60,5
	1	391	587	60
Overall Percentage				60,3

a. The cut value is ,500

Quadro 38 - Matriz de Classificação – Modelo “Clientes VN < 2 Milhões Euros”

		Predicted		
		Y		Percentage Correct
		0	1	
Observed				
Step 4 Y	0	206	305	40,3
	1	128	467	78,5
Overall Percentage				60,8

a. The cut value is ,500

Quadro 39 - Matriz de Classificação – Modelo “Clientes VN > 2 Milhões Euros”

		Predicted		
		Y		Percentage Correct
		0	1	
Observed				
Step 8 Y	0	407	73	84,8
	1	251	124	33,1
Overall Percentage				62,1

a. The cut value is ,500

Quadro 40 – Matriz de Classificação – Modelo “Clientes VN < 10 Milhões Euros”

		Predicted		
		Y		Percentage Correct
		0	1	
Observed				
Step 6 Y	0	460	348	56,9
	1	295	526	64,1
Overall Percentage				60,5

a. The cut value is ,500

Quadro 41 - Matriz de Classificação – Modelo “Clientes VN > 10 Milhões Euros”

			Predicted		
			Y		Percentage Correct
			0	1	
Step 2	Observed				
	Y		0	1	
		0	143	35	80,3
		1	95	53	35,8
	Overall Percentage				60,1

a. The cut value is ,500

Quadro 42 – Matriz de Classificação – Modelo “Clientes VN < 50 Milhões Euros”

Observed			Predicted		
			Y		Percentage Correct
			0	1	
Step 7	Y	0	586	352	62,5
		1	384	544	58,6
Overall Percentage					60,6

a. The cut value is ,500

Quadro 43 - Matriz de Classificação – Modelo “Clientes VN > 50 Milhões Euros”

Observed			Predicted		
			Y		Percentage Correct
			0	1	
Step 1	Y	0	41	7	85,4
		1	29	12	29,3
Overall Percentage					59,6

a. The cut value is ,500

O texto foi escrito ao abrigo do Novo Acordo Ortográfico